Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 4, Issue 5, 2026

# A COMPREHENSIVE REVIEW OF DEEP LEARNING ADVANCEMENTS IN EDUCATION: CHALLENGES AND FUTURE DIRECTIONS

**Mashal Tariq[*1], Shehla Andleeb[2], Rabbia Muhammad Qasmi[3], Durr-e-Shahwar[4], Hafsa Asif[5], Muhammad Shakir Khan[6]**

[*1,2]*Senior Lecturer DHA Suffa University, Karachi, Pakistan.*
[3]*Lecturer DHA Suffa University, Karachi, Pakistan.*
[4,5,6]*Junior Lecturer DHA Suffa University, Karachi, Pakistan.*

[*1]mashal.tariq@dsu.edu.pk,[2]shehla.andleeb@dsu.edu.pk,[3]rabbia.qasmi@dsu.edu.pk,
[4]durre.shehwar@dsu.edu.pk, [5]hafsa.asif@dsu.edu.pk , [6]shakir.khan@dsu.edu.pk

**DOI:** https://doi.org/10.5281/zenodo.20268474

**Keywords**
*Deep Learning; Artificial Intelligence; Education; Personalized Learning; Student Performance Prediction; Emotion Detection; Educational Data Mining.*

**Article History**
*Received: 11 March 2026*
*Accepted: 21 April 2026*
*Published: 13 May 2026*



**Corresponding Author:** *
**Mashal Tariq**

**Abstract**
*Deep learning (DL) has emerged as a transformative paradigm in education, enabling intelligent, automated, and data-driven solutions across multiple dimensions of teaching and learning. This review provides a comprehensive analysis of recent advancements in DL applications within education, focusing on personalized learning, automated evaluation, student performance prediction, sentiment analysis, and learning engagement. The paper highlights how DL models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), and hybrid architectures are being employed to enhance teaching efficiency, optimize learning outcomes, and improve decision-making in educational institutions. Key contributions include the identification of novel frameworks for real-time monitoring, emotion detection, and cybersecure access to learning platforms. The review also examines major challenges such as dataset scarcity, lack of scalability, ethical concerns, and limitations in real-world integration. Finally, future research directions are outlined, emphasizing the development of unified, holistic, and ethically responsible DL-powered education systems that integrate academic, behavioral, and administrative functions.*

## 1. INTRODUCTION

The rapid integration of artificial intelligence (AI) into education has transformed traditional teaching and learning practices into more adaptive, personalized, and data-driven processes [1]. Among various AI techniques, deep learning (DL) has attracted increasing attention due to its ability to process complex, multimodal data such as text, speech, video, and behavioral patterns with high accuracy and scalability [2]. Unlike conventional rule-based or statistical approaches, DL models can automatically extract meaningful features from large datasets, enabling intelligent decision-making and predictive analytics in education.

Over the past decade, DL has been applied in multiple areas of education including automated assessment, personalized learning, student performance prediction, engagement monitoring, literacy management, and cybersecurity in online learning systems [3]. The adoption of these technologies accelerated further during the COVID-19 pandemic, which highlighted the need for scalable and intelligent educational platforms capable of supporting remote and hybrid learning environments. Models such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, Graph Neural Networks (GNNs), and Transformer-based architectures have shown significant promise in improving evaluation accuracy, identifying at-risk students, detecting emotions, and supporting decision-making for teachers and administrators.

Several review studies have explored artificial intelligence and machine learning in education, primarily focusing on educational data mining, traditional neural networks, or isolated applications such as student performance prediction. For example, Okewu et al. (2021) provided a large-scale survey of educational data mining techniques, while Alnasyan et al. (2024) reviewed DL applications in MOOCs, and Yousafzai et al. (2021) examined Bi-LSTM models for performance prediction. While these works provide valuable insights, they are limited either in scope, methodology, or coverage of recent developments. Most do not address the broader integration of DL across multiple educational dimensions, nor do they sufficiently discuss limitations such as dataset scarcity, scalability, ethical implications, and the lack of unified real-time systems.

This paper aims to fill that gap by presenting a comprehensive and up-to-date review of DL applications in education. It not only synthesizes existing research across diverse application areas but also critically evaluates methodologies, datasets, and performance outcomes. In doing so, the paper contributes threefold:

1. It categorizes DL applications into key domains such as evaluation, engagement, prediction, sentiment analysis, literacy management, and cybersecurity.
2. It highlights open challenges including data imbalance, lack of scalability, and ethical issues that hinder real-world deployment.
3. It outlines future directions, proposing a roadmap toward the development of unified, ethical, and holistic DL-powered education systems.

By consolidating recent literature and identifying research gaps, this review offers a valuable reference for researchers, educators, and policymakers seeking to harness the potential of DL to improve teaching, learning, and educational management

## 2. Review of Recent Research in Education

### 2.1 Evaluation and Teaching Assessment

Evaluation is a cornerstone of education, and the automation of this process using deep learning (DL) has significantly benefited teachers and institutions. A Convolutional Neural Network (CNN) model [4] was employed at the university level to address challenges in traditional evaluation systems, such as subjectivity and lack of real-time feedback. The model classified outcomes into four categories—Excellent, Good, Average, and Poor—with an accuracy ranging between 80–92% across classifiers. Inputs included multi-source data such as student performance, teacher evaluation, student feedback, course content, and behavioral data. Teacher performance was measured using three critical features: content clarity, classroom interaction, and teacher attitude.

Subject-specific teaching evaluation has also been explored. For example, the Genetic Algorithm-Backpropagation Neural Network (GA-BPNN) [5] was applied in Physical Education assessments. This hierarchical evaluation method outperformed traditional approaches, demonstrating the adaptability of DL in diverse academic contexts.

### 2.2 Social Media and Technology Impact on Learning

DL models have also been employed to study the influence of social media and technology usage on student learning. A Gated Attention DL Model [6], trained on data from the Kalboard 360 LMS (available on Kaggle), categorized student performance based on technology interaction. The analysis revealed that students effectively using technology achieved better outcomes, while excessive social media usage correlated with lower performance. Furthermore, parental education, particularly that of mothers, emerged as a significant factor in student achievement. The model attained 98–99% accuracy and precision, underscoring the strong predictive power of DL in this domain.

### 2.3 Behavioral Analysis and Classroom Monitoring

Habitual behaviors that influence academic performance have also been studied using DL approaches. Mobile phone usage in classrooms, for instance, was detected using a ResNet50 architecture integrated with an RTL model-based approach [7]. The system achieved 96% accuracy in identifying students' phone usage, though its lack of a defined long-term goal and real-time deployment strategy highlights areas for future research.

Similarly, predicting student performance has been tackled using advanced hybrid models. The GNN-TINet (Graph Neural Network–Transformer–InceptionNet) [8] achieved 98.5% accuracy in linking GPA with parental involvement in homework guidance. Two novel indicators—Learning Impact Factor (LIF) and Predictive Consistency Score (PCS)—were introduced, adding to existing performance metrics. Despite promising outcomes, reliance on publicly available statistics raised concerns about applicability across diverse educational contexts.

### 2.4 Literacy Feedback and Performance Prediction

Another stream of research emphasizes literacy and retention. A literacy management system [9] allowed students to log in, take tests, and receive personalized literacy rates. Performance improved when students actively engaged with the system.

To address challenges of student retention and admissions, DL models have been applied in education data mining (EDM). A study comparing multiple performance prediction models used Conditional Tabular Generative Adversarial Networks (CTGAN) against Gaussian distribution. CTGAN provided better results but showed limitations in balancing datasets, as reflected in lower F1-scores [10]. Furthermore, a systematic review of ANN and DL applications in EDM (2010–2018) highlighted the breadth of models but lacked quantitative details [11].

### 2.5 Emotion Detection and Online Learning

The growth of online education has motivated research into emotion and sentiment analysis. Using the Spatial Attention Network with CNN (SAN-CNN) [12], one study achieved 98% accuracy in detecting student emotions during online classes. While effective for recognition, integration into real-world learning environments remains untested.

Similarly, cybersecurity in higher education has become crucial, particularly with the rise of online teaching during COVID-19. The Hunter-Prey Optimizer with DL-enabled Biometric Verification for Cybersecurity (HPODL-BVCS) [13] was developed to protect educational resources. Incorporating eye and handprint biometrics, the system achieved 99% accuracy in preventing unauthorized access.

Xing et al. (2025) [15] introduced a smart education system using a simple neural network to assess student performance with multimodal data (voice, text, and image). The system employed random weight initialization and backpropagation for optimization, though results were not fully reported, leaving a gap in evaluating its effectiveness.

### 2.6 MOOCs and Student Engagement

Massive Open Online Courses (MOOCs) have expanded global access to training and skill development. A review of 46 studies [16] demonstrated how DL, in combination with ML techniques, enhances student performance prediction in MOOCs. Both qualitative and quantitative insights confirmed that success in MOOCs increasingly depends on practical skills rather than academic credentials.

Offline and online engagement has also been studied through behavioral tracking. IoT-based frameworks have been proposed to analyze student behavior using eye movements, facial expressions, and head gestures [17]. However, reliance on IoT makes real-time deployment in classrooms impractical. Alternatively, a 3D CNN model [18] achieved ~99% accuracy in evaluating online student engagement, showing high promise for scalable solutions.

Despite these advancements, existing research has notable limitations, which we discuss below. The aforementioned research studies are from the year 2025, while additional research details are provided in the table. Considering the length and goal of this publication, the primary aim of this paper is to present various research areas in education along with their DL models, approaches, and results. Table 1 summarizes the most recent literature review and new models with its applications, limitations etc.

**Table 1 Summary of Literature Review**

| *Reference* | *Study Focus* | *DL Model* | *Dataset Details* | *Limitation* | *Future Gap* | *Key Findings* |
|---|---|---|---|---|---|---|
| *[4] Gao (2025)* | University teaching evaluation | CNN | Multi-source demographic data (student performance, teacher evaluation, etc.) | Versatility of dataset and student sentiments are not addressed | Large dataset. Inclusion of student's emotions. Integration of ML models. | Accuracy: 80-92% |
| *[5] Chen & Min (2025)* | Physical education teaching performance | GA-BPNN | Student feedback, peer reviews and self assessment reports | Details of dataset are missing. Performance indicators are missing | Quantitative Research | Better results than industry average |
| *[6] Kumar et al. (2025)* | Social media influence on education | A-Bi-LSTM-CNN-G Bi-LSTM-CNN Bi-LSTM CNN LSTM RNN | Kalboard 360 LMS | The size of data and processing time | Other technological Factors to be included | Accuracy: 96-99% |
| *[7] Pandya et al. (2025)* | Mobile phone usage detection | ResNet50 RTL Model-Based approach | ImageNet dataset (Deng et al., 2009) for training where for testing real CCTV footage of 20 students [14] | Small testing dataset | Large dataset testing | Accuracy: 96% |
| *[8] Zhang et al. (2025)* | Student performance prediction | GNN-TINet | Publicly available (California Schools) GPA, | Demographic | Testing and training on real data set | Accuracy: 98.5% |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | parental involvement | | | |
| *[9] Zhang (2025)* | Literacy management system | DL algorithms | Student literacy data. Online | Small Sample Size | Inclusion of user satisfaction and system optimization | Improved student involvement |
| *[10] Kannan et al. (2025)* | Education data mining | CTGAN | Real dataset | Data level techniques. Small data set | Adding more algorithm at different levels. Large dataset. More DL synthesizers | F1 scores low due to class imbalance |
| *[11] Okewu et. Al (2021)* | Education data mining | ANN (DL) | Review Paper.577 (Research Studies) | Quantitative of research studies | EDM Expert System | 73 out 577 shows DL research |
| *[12]Mahendar et al. (2025)* | Emotion detection in online learning | SAN-CNN | Online (Kaggle). Cognitive state data | Real time Data | More emotion and enhancement of technique | Accuracy: 98% |
| *[13] Mary & Claral (2025)* | Student sentiment analysis | Hybrid DL Model NLP, CNN & LSTM | Real world data. Sentiment data | Dedicated to education field and the student feedback | The same technique to be applied on another field like brand management | High accuracy in sentiment prediction |
| *[15] Xing et al. (2025)* | Intelligent education evaluation | Neural Network | Details of dataset are missing. Voice, text, image data | Proposed Idea implementation is missing | Qualitative and Quantitative data. | No results provided |
| *[16] Alnasyan et al. (2024)* | Student Performance in MOOC | DNN, LSTM, CNN | Review Paper 46 studies | Data Availability Performance Parameter Hybrid Models Predicting the performance at the beginning and end of | Hybrid and Comprehensive complete Predictive Performance Model | Accuracy 90% |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | the semester is missing | | |
| *[17] Yousafzai, B. K. et. Al (2021).* | Student Performance | BiLSTM | UCI ML Repository [18 ] | Versatility of dataset | Real time dataset and integration of other DL models | Prediction accuracy of 90.16% |
| *[19] Akila, D. et.al (2024).* | Offline Student Engagement Evaluator | DL-student attention recognition framework (DL-SARF) NLP | Online | During the class it is hard to keep track of the student's attention This could be only used to give indications IoT involvement | Make simpler model and more efficient model. | High student performance ratio of 98.2 %, learning outcome ratio of 97.0%, enhanced prediction ratio of 98.2 %,feedback ratio of 97.5 %, accuracy ratio of 98.4 %, behavior rate of 97.4 %, interaction ratio of 97.9 % |
| *[20] Mehta, N. K. et.al (2022).* | Online Student Engagement Evaluator | DenseNet self-attention neural network (DenseAttNet) u | DAiSEE and EmotiW-EP datasets online | Lower Accuracy | More DL models for multi class and multi label advanced data sets. | DenseAttNet trained labels; namely boredom, engagement, confusion, and frustration has an accuracy of 81.17%, 94.85%, 90.96%, |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | and 95.85% |
| *[21] Yu, H.(2023)* | Education Sector | Generative AI-DL | - | Qualitative and Quantitative Data | Qualitative and Quantitative Data | Discussion |

**2.2 Future Directions**

Future research should focus on integrating various DL models into a unified system, ensuring scalability with real-time data, and addressing demographic and behavioral variations in student populations. Additionally, ethical considerations, such as data privacy and fairness in AI-driven educational assessments, should be explored.

There are many systematic papers found during the research, covering ML, DL, and artificial neural networks. However, since the focus here is on DL, preference is given to artificial intelligence-driven DL models to ensure relevant and valuable material. Some of these papers are included, whereas others are repetitions of previously discussed applications.

The review could be really helpful for the students, while the main purpose of this paper is to present the maximum area of recent studies, so the focus is to include different dimensions in the particular sector.

For research using real-time data, the size of the dataset remains a major limitation. Models trained on small or medium-sized datasets cannot always be considered reliable for large-scale applications. Addressing this gap could improve models for behavioral analysis, dropout prediction, and real-time learning analytics.

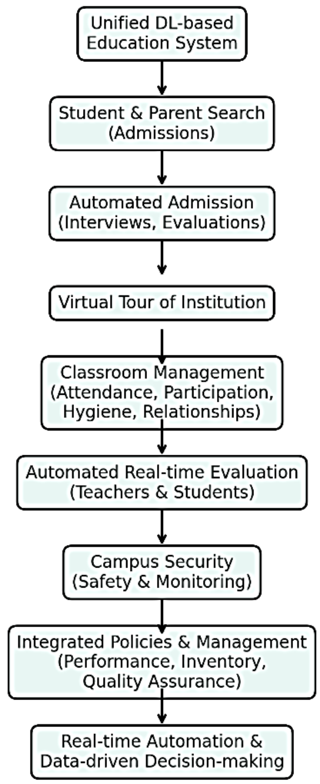
Moreover, the need of a single comprehensive research paper that presents a fully integrated DL-based education system still exists. Most studies focus on specific aspects of education rather than providing a holistic approach. For example, an ideal system would integrate:

- Student and parent search functionalities for admissions
- Automated admission processes, including interviews and evaluations
- A virtual tour of the institution
- Classroom management, covering attendance, participation, hygiene, and teacher-student relationships
- Automated real-time evaluation for both teachers and students
- Campus security, ensuring safety in classrooms and surrounding areas
- Comprehensive integration of education policies, performance tracking, inventory management, and quality assurance

For such a system to be effective, it must be capable of real-time automation and data-driven decision-making while ensuring security and privacy. A unified, AI-powered software platform should integrate all these aspects to enhance the efficiency and effectiveness of modern education systems.

Figure 1. shows a proposed DL-based Education System. This figure shows the proposed system for future educational frameworks, illustrating how deep learning can integrate multiple components—ranging from admissions to classroom management, evaluation, security, and real-time automation—into a unified platform.

**Linear Flow of DL-based Education System**




## 3. Conclusion

Deep learning has demonstrated significant potential to revolutionize education by enabling intelligent, personalized, and automated systems that benefit students, teachers, and administrators alike. From evaluation and performance prediction to emotion detection and cybersecurity, DL applications have expanded across diverse aspects of educational ecosystems. This review highlights the strengths of recent DL models while also pointing out the recurring challenges of dataset limitations, model scalability, ethical concerns, and real-time applicability.

The findings suggest that while DL has made considerable progress in isolated tasks, the next step lies in developing comprehensive and unified frameworks that integrate evaluation, engagement, security, and administrative functions into a single intelligent platform. Future research must also prioritize fairness, transparency, and privacy to ensure that DL-driven systems enhance learning outcomes without reinforcing existing biases.

In conclusion, deep learning holds the promise of reshaping education by bridging gaps between traditional pedagogy and modern technological demands. A collaborative effort from researchers, policymakers, and educational institutions will be crucial in designing scalable, ethical, and holistic DL-powered systems capable of meeting the evolving needs of global education.

**Acknowledgement**

The authors acknowledge the resources provided by DHA Suffa University and Dawood University of Engineering and Technology for the collection of review articles throughout the research and writing process.

We are also deeply grateful to our families for their unwavering support, patience, and motivation, which enabled us to successfully complete this paper.

**Declaration of Interest Statement**

The authors confirm that there are no financial or personal affiliations or conflicts that could have affected the objectivity or integrity of the research presented in this paper.

**REFERENCES:**

[1] Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2025). AI-driven adaptive learning for sustainable educational transformation. *Sustainable Development*, *33*(2), 1921-1947.

[2] Summaira, J., Li, X., Shoib, A. M., Li, S., & Abdul, J. (2021). Recent advances and trends in multimodal deep learning: A review. *arXiv preprint arXiv:2105.11087*.

[3] Sengupta, S., & Chakrabarti, S. (2025). Towards a smarter education system: an investigation into ML and DL for information retrieval. *Multimedia Tools and Applications*, 1-34.

[4] Gao, Y. (2025). DL-based strategies for evaluating and enhancing university teaching quality. *Computers and Education: Artificial Intelligence*, *8*, 100362.

[5] Chen, W., & Min, L. (2025). The Evaluation Model of Physical Education Teaching Performance based on DL Algorithm. Scalable Computing: Practice and Experience, 26(1), 458-466.

[6] Praveen Kumar, B., Kalpana, A. V., & Nalini, S. (2025). Gated Attention Based DL Model for Analyzing the Influence of Social Media on Education. *Journal of Experimental & Theoretical Artificial Intelligence*, *37*(2), 291-305.

[7] Pandya, D., Rana, K., & Padhiyar, A. (2025). Comparative analysis of DL approach to identify usage of mobile phone in the classroom environment. *Education and Information Technologies*, 1-24.

[8] Zhang, X., Zhang, Y., Chen, A. L., Yu, M., & Zhang, L. (2025). Optimizing multi label student performance prediction with GNN-TINet: A contextual multidimensional DL framework. *PloS one*, *20*(1), e0314823.

[9] Z hang, D. (2025). Design and Implementation of an Information Literacy Enhancement System for College Students Based on DL Algorithms. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, *20*(1), 1-20.

[10] Kannan, R., Ng, H., Yap, T. T. V., Wong, L. K., Chua, F. F., Goh, V. T., ... & Wong, H. L. (2025). Handling class imbalance in education using data-level and DL methods. *International Journal of Electrical and Computer Engineering*, *15*(1), 741-754.

[11] Okewu, E., Adewole, P., Misra, S., Maskeliunas, R., & Damasevicius, R. (2021). Artificial neural networks for educational data mining in higher education: A systematic literature review. *Applied Artificial Intelligence*, *35*(13), 983-1021.

[12] Mahendar, M., Malik, A., & Batra, I. (2025). Emotion estimation model for cognitive state analysis of learners in online education using DL. *Expert Systems*, *42*(1), e13289.

[13] Mary, T., & Claral, A. (2025). Hybrid DL Model to Predict Students' Sentiments in Higher Educational Institutions. *International Journal of Interactive Mobile Technologies*, *19*(1).

[14] Xing, Z., Yang, Y., Tan, L., & Guo, X. (2025). Multi-source physical information driven DL in intelligent education: Unleashing the potential of deep neural networks in complex educational evaluation. *AIP Advances*, *15*(2).

[15] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, Li. (2009). ImageNet: A large-scale hierarchical image database. IEEE Conference on Computer Visionand Pattern Recognition, 2009, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

[16] Alnasyan, B., Basheri, M., & Alassafi, M. (2024). The power of DL techniques for predicting student performance in Virtual Learning Environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 100231.

[17] Yousafzai, B. K., Khan, S. A., Rahman, T., Khan, I., Ullah, I., Ur Rehman, A., ... & Cheikhrouhou, O. (2021). Student-performulator: student academic performance using hybrid deep neural network. *Sustainability*, *13*(17), 9775.

[18] Musleh, M., Ouzzani, M., Tang, N., & Doan, A. (2020, June). Coclean: Collaborative data cleaning. In *Proceedings of the 2020 ACM SIGMOD international conference on management of data* (pp. 2757-2760).

[19] Akila, D., Garg, H., Pal, S., & Jeyalaksshmi, S. (2024). Research on recognition of students attention in offline classroom-based on DL. *Education and Information Technologies*, *29*(6), 6865-6893.

[20] Mehta, N. K., Prasad, S. S., Saurav, S., Saini, R., & Singh, S. (2022). Three-dimensional DenseNet self-attention neural network for automatic detection of student's engagement. *Applied Intelligence*, *52*(12), 13803-13823.

[21] Yu, H., & Guo, Y. (2023, June). Generative artificial intelligence empowers educational reform: current status, issues, and prospects. In *Frontiers in Education* (Vol. 8, p. 1183162). Frontiers Media SA.