Data Engineering and Applications

Review Paper

Data-Driven Approaches to Behavior Analysis in Smart Classrooms: A Review of AI Algorithms

Damian Wisniewski[1,*], Marek Zielinski[1], Patrycja Dabrowski[2]

[1] Faculty of Computer Science and Information Technology, Warsaw University of Technology, Warsaw 00-665, Poland
[2] Faculty of Computer Science and Management, Silesian University of Technology, Gliwice 44-100, Poland
*Corresponding author: d.wisniewski@pw.edu.pl

Abstract. With the rapid advancement of educational informatization, smart classrooms—as new teaching ecosystems integrating multiple sensing devices and data streams—are profoundly transforming the observation and analysis of classroom behaviors. Artificial intelligence algorithms, particularly machine learning and deep learning methods, provide a solid foundation for the efficient mining and intelligent interpretation of behavioral data, helping to reveal students' learning states, interaction patterns, and cognitive dynamics. This paper systematically reviews core domestic and international literature from 2015 to 2022, focusing on artificial intelligence algorithms in the field of intelligent classroom behavior analysis. The research encompasses key technical approaches including single-modal, multi-modal fusion, deep learning, and edge computing. By comparing multiple dimensions including recognition accuracy, robustness, real-time performance, and scalability, the study further highlights the significant advantages of multimodal fusion and deep learning architectures in complex behavior recognition. It simultaneously identifies key challenges such as data scarcity, insufficient model generalization capabilities, interpretability, and privacy protection. This paper summarizes the urgent need to establish a standardized evaluation system and explores the latest trends in human-machine collaboration. The conclusion emphasizes that future advancements in intelligent classroom behavior analysis must rely on interdisciplinary innovation to drive algorithmic interpretability, privacy compliance, and deep integration with real teaching scenarios. This will facilitate the transition of behavioral intelligence analysis from experimental validation to large-scale application. This research aims to provide systematic theoretical guidance and a technical roadmap for scholars and practitioners in educational informatization, promoting the standardization and sustainable development of intelligent classroom behavior analysis research.

Keywords: *Computer Vision, Multimodal Fusion, Behavior Analysis, Smart Classroom, Artificial Intelligence*

Received on 19 August 2024, Accepted on 3 November 2024, Published on 11 December 2024

Introduction

The development of smart classroom technology has revolutionized contemporary education. It has elevated interactivity, adaptability, and data centralization to unprecedented levels. The smart classroom is not merely the simple application of digital whiteboards or networked devices; it is an ecosystem composed of wearables, cameras, sensors, and heterogeneous data streams, enabling deeper observation and analysis of classroom dynamics [1]. Within this environment, behavioral analytics has become foundational to educational theory and practice. By collecting and analyzing student behaviors—ranging from postures and gestures to patterns of interactive engagement—researchers gain insights into learning processes, cognitive states, and social dynamics. Traditional classroom behavior observation relied on time-consuming manual coding or self-reporting, both methods suffering from high subjectivity and scalability limitations. The application of artificial intelligence (AI) in education has revolutionized the possibility of objective, scalable, and real-time behavioral analysis. With the rise of computer vision and sensor fusion technologies, machine learning systems can now autonomously identify subtle behavioral cues such as inattentiveness, collaborative interactions, states of confusion, or emotional fluctuations. The accuracy and consistency of detecting these nuanced behavioral cues often surpass human observation [2][3].This technological transformation aligns with the broader trend of data-driven teaching, where continuous multimodal data streams inform instructional decisions, personalize learning paths, optimize classroom management, and enhance student learning quality [4]. Most significantly, AI-driven behavioral analytics transcends mere observation automation. It translates concepts from educational psychology and learning sciences—such as engagement, motivation, and cognitive load—into quantifiable metrics for systematic monitoring and actionable insights [5]. It offers two key benefits. First, it enables formative assessment and adaptive feedback with unprecedented granularity. Second, it supports the empirical validation of instructional interventions in real-world settings. Nevertheless, this prospect remains fraught with challenges. Discussions are ongoing within computer science and education communities regarding the epistemological validity of algorithmic reasoning, the risks of algorithmic bias, and the ethical implications of pervasive monitoring [6].

Given the rapid advancement of recent research and its interdisciplinary nature, a critical and comprehensive evaluation of artificial intelligence algorithms in intelligent classroom behavior analysis is both timely and necessary. This review primarily discusses computational methods, particularly machine learning, deep learning, and multimodal data fusion techniques, which will be validated in the educational field between 2015 and 2022. It covers foundational approaches and recent advancements through selected peer-reviewed journal articles, flagship conference proceedings, and influential review literature [7][8].This review aims to achieve four core objectives: categorize AI algorithms for intelligent classroom behavior analysis; compare the strengths and weaknesses of various algorithm types; integrate emerging trends and unresolved issues; and contrast the current state of research with practical application scenarios. This review excludes studies on algorithmic behavior analysis, hardware-oriented classroom technologies, or research focusing solely on teacher behavior without addressing student interaction analysis. The literature selection process prioritized methodological innovation, empirical rigor, and relevance to smart classrooms, though it did not employ a fully systematic review methodology. Its purpose extends beyond mere summarization; it also commits to critically evaluating and analyzing key research findings to provide a more comprehensive perspective on the current research landscape.

The purpose of this discussion is to progressively build a comprehensive understanding of this field. The next section will introduce a classification system for artificial intelligence algorithms in intelligent classroom behavior analysis. It will distinguish between unimodal approaches, multimodal fusion strategies, and major deep learning model families. It will also discuss edge AI and real-time system implementation solutions. Subsequently, a comparative analysis of these algorithms will be conducted to evaluate key performance metrics—such as accuracy, robustness, latency, and scalability—along with application scenarios including engagement detection and anomaly identification. To provide practitioners and researchers with a clear reference benchmark, tables will systematically summarize and contrast algorithmic characteristics and empirical results. In subsequent sections, we will discuss data annotation challenges, model transferability, and algorithmic bias. These issues are comprehensively analyzed by examining current trends, unresolved questions, and theoretical debates. The findings include research supporting and challenging contemporary approaches. The review concludes by outlining future research directions, including explainable artificial intelligence, privacy-preserving machine learning, and benchmark and dataset standardization. It also identifies potential obstacles that may be encountered.

Classification of AI Algorithms for Behavior Analysis

The classification system for AI algorithms in intelligent classroom behavior analysis is crucial, as it helps elucidate their scope of application, applicability, and methodological development. Edge-oriented real-time deployment, deep learning model architectures, multimodal fusion frameworks, and unimodal algorithms represent the four core directions of contemporary research. As shown in Figure 1, the core artificial intelligence algorithms for intelligent classroom behavior analysis are categorized into four major types: edge AI, deep learning, multimodal fusion, and unimodal.

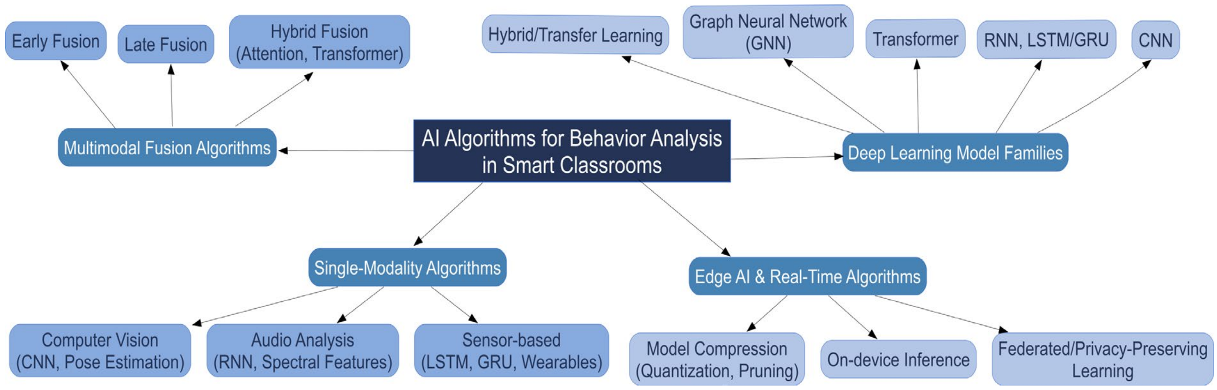

Figure 1 Taxonomy of AI Algorithms for Behavior Analysis in Smart Classrooms

Single-Modality Algorithms

Unimodal methods are crucial for computational modeling of classroom behavior, offering clear and easily interpretable signal processing. Computer vision-based algorithms, particularly those utilizing convolutional neural networks (CNNs), demonstrate outstanding performance in tasks such as gesture analysis, body posture recognition, and facial emotion recognition [11]. Spatial feature extractors based on neural networks (CNNs) are typically combined with region proposal mechanisms and attention maps to better distinguish complex classroom behaviors [12]. Pose estimation architectures like OpenPose and AlphaPose capture skeletal dynamics through local affinity fields and multi-stage refinement networks, enabling more precise analysis of engagement and limb interactions [13]. The expressiveness of pose models is particularly vital in collaborative learning. Spatial distance and gesture synchrony serve as key metrics for evaluating group interaction quality [14].

Acoustic and audio-based algorithms represent another parallel development path. These algorithms can utilize both time-domain and frequency-domain characteristics to identify speech, silence, and paralinguistic cues [15]. By employing Mel-frequency cepstral coefficients (MFCCs), spectrogram analysis, and deep recurrent neural networks, pedagogically significant events such as hesitation phenomena, vocal engagement, and classroom discourse segmentation can be detected [16]. Sensor-based solutions can covertly identify subtle movements, postural shifts, and device interactions through data from accelerometers, gyroscopes, and pressure-sensing wearables [17]. These algorithms are particularly adept at capturing behaviors that evade visual or auditory monitoring due to occlusion or environmental noise.

Despite the numerous advantages of unimodal systems, they remain susceptible to risks such as modality-specific noise, domain adaptation challenges, and algorithmic blind spots [18]. Despite these limitations, such algorithms remain indispensable, as their computational efficiency and interpretability render them highly valuable in scenarios with limited infrastructure or stringent latency requirements.

Multimodal Fusion Algorithms

Theoretical and practical advancements in multimodal fusion algorithms have redefined the upper limits of behavioral analysis precision and ecological validity. Early fusion strategies concatenated visual, auditory, and physiological feature vectors prior to model input, but this sometimes resulted in reduced modality-specific resolution [19]. In contrast, late-stage fusion frameworks employ stacking, weighted voting, or Bayesian integration to aggregate modality-specific inference streams into final predictions. This approach enhances robustness against missing or corrupted channels while preserving interpretability [20]. Attention-based gating mechanisms combine early and late fusion by dynamically adjusting modality weights based on signal quality or environmental changes. Consequently, hybrid fusion frameworks have emerged [21].

Deep multimodal learning models, particularly those integrating Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN) for spatial encoding and temporal sequence modeling, have achieved state-of-the-art results in studies such as object recognition and collaborative activity detection [22]. Transformer-based architectures have further advanced this field by enabling long-range dependency modeling and adaptive contextual integration through self-attention mechanisms and cross-modal alignment techniques [23]. These models can fuse sensor log information, audio traces, and video streams to achieve high-dimensional embeddings that preserve modality specificity and cross-modal dynamics.

Empirical studies demonstrate that multimodal fusion frameworks exhibit significant advantages in scenarios involving complex, ambiguous, or transient behaviors [24]. These scenarios include group problem-solving and unstructured discussions. However, such systems require complex alignment algorithms, precise synchronization of heterogeneous data streams, and substantial computational resources, making deployment in resource-constrained environments challenging. Striking a balance between multimodal comprehensiveness and operational scalability remains a primary challenge in both technology and practice. As illustrated in Figure 2, the overall workflow of multimodal fusion for classroom behavior analysis typically comprises three stages: feature extraction, fusion mechanisms, and behavior recognition.

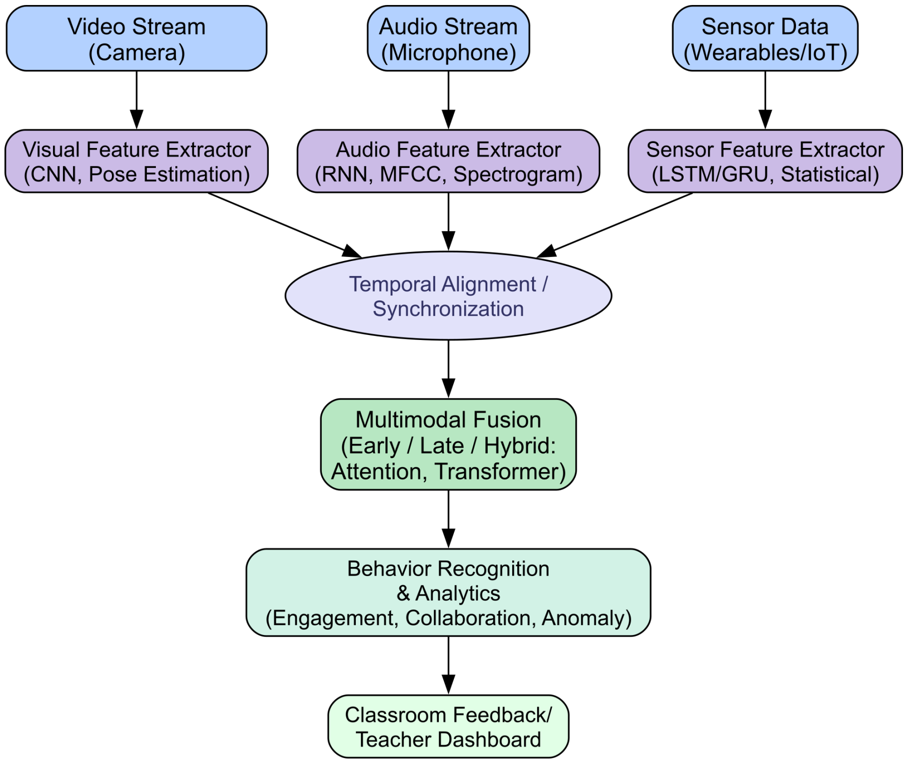

Figure 2 Typical Multimodal Fusion Framework for Behavior Analysis

Deep Learning Model Families

Many recent advancements in intelligent classroom behavior analysis have benefited from the maturation of deep learning architectures. Convolutional neural networks (CNNs) excel at extracting spatial features from image and video data by employing hierarchical, locally sensitive filters [11][12]. To capture temporal dependencies in audio, sensor, and multimodal data streams, recurrent neural networks (RNNs) and their variants—such as LSTM and gated recurrent unit (GRU) models—are widely applied in time-series event modeling [16][22]. Transformer models, based on self-attention mechanisms and parallel processing capabilities, can handle long-range dependencies and complex cross-modal interactions. They excel in behavioral analysis tasks spanning extended time periods or rich contextual environments [23].

Graph neural networks (GNNs) have been applied to relational reasoning, enabling the modeling of social networks within classrooms and facilitating analysis of collaborative dynamics, peer influence, and interaction topologies [25]. Hybrid models combining spatial (convolutional neural networks) and temporal (long short-term memory networks, transformers) modules enhance representational capabilities, enabling joint modeling of dynamic and static behavioral features. Many researchers are employing transfer learning and domain adaptation methods to address data scarcity and domain shift. These approaches first pre-train models on large-scale speech or vision corpora, then fine-tune them on specific classroom datasets [14][26].

Deep learning models have redefined the technological frontier, yet their application in education is constrained by challenges such as interpretability, data scarcity, and the opacity of learning representations. Due to the black-box nature of deep architectures, algorithmic outputs prove difficult to translate into actionable pedagogical insights. Consequently, practical deployment necessitates the adoption of interpretable model transparency solutions and AI explanations.

Edge AI and Real-Time Algorithms

The emergence of edge AI and real-time inference marks a pivotal turning point in the operationalization of behavioral analytics systems. Deploying lightweight deep learning models on embedded GPUs, single-board computers, and smartphones effectively reduces latency, enhances privacy protection, and improves system resilience by eliminating the need for persistent cloud connectivity [17][18]. Deploying convolutional neural networks (CNNs) and recurrent neural networks (RNNs) within computational and power constraints necessitates model compression techniques such as pruning, quantization, and knowledge distillation [13][26]. These optimizations simultaneously reduce computational cycles and memory usage while maintaining widely accepted inference accuracy.

Edge-oriented systems are particularly crucial in educational settings with inadequate or unreliable network infrastructure, or when data privacy regulations prohibit transmitting raw sensor or video data to external servers [20]. Real-time behavioral analysis facilitates collaborative classroom management, adaptive feedback, and immediate instructional interventions. This aligns with theoretical frameworks for formative assessment and instant learning support [15]. However, integrating edge AI introduces new engineering challenges, such as model updates, federated learning for decentralized data aggregation, and device-side anomaly detection. Table 1 provides a detailed taxonomy of representative AI algorithms for classroom behavior analysis, mapping each to its data modality, model family, principal application, and key advantages and limitations.

Table 1. Taxonomy of AI Algorithms for Smart Classroom Behavior Analysis: Modalities, Models, Applications, and Technical Considerations

Modality	Model Family	Representative Application	Key Advantages	Primary Limitations
Vision	CNN, Pose Estimation	Posture gesture, facial affect	High spatial fidelity, direct mapping to observable behavior	Sensitive to occlusion lighting variation, privacy concerns
Audio	RNN CNN	Speech activity engagement, emotion	Temporal granularity, unobtrusive	Susceptible to noise ,limited in silent behaviors
Sensor	LSTM, GRU, DNN	Movement analytics, device usage	Fine-grained micro-movement capture, minimal visual intrusion	Limited context, hardware dependence
Multimodal Fusion	CNN + LSTM, Transformer	Collaboration group dynamics, engagement	Robustness contextual richness, cross-validation	High resource demand, complex synchronization
Edge AI	Quantized CNN, Pruned RNN	Real-time inference, privacy-preserving analytics	Low latency local processing, regulatory compliance	Model accuracy trade-off, update complexity
GNN	Graph Neural Network	Social network analysis, peer influence	Relational modeling ,interaction topology	Requires labeled interaction graphs ,interpretability issues

Comparative Analysis of Methods

Comparative analysis of artificial intelligence algorithms for intelligent classroom behavior analysis requires careful examination across multiple dimensions, including datasets, metric systems, algorithmic strengths and weaknesses, and scenario-dependent performance. Critical integration and comprehensive benchmarking are essential for guiding theoretical development and practical application.

Evaluation Metrics and Datasets

The multi-indicator framework serves as the scientific benchmark for intelligent classroom behavior analysis. Accuracy remains a critical metric. Quantifying algorithms through precision, recall, F1 score, and area under the curve (AUC) establishes their correspondence with labeled ground truth [27]. Real-time intervention and adaptive feedback systems require latency—the time interval between data acquisition and inference output [28]. Robustness can be evaluated by an algorithm's performance under conditions of data corruption, occlusion, sensor noise, and domain transfer. Recently, this characteristic has been measured through cross-dataset validation and adversarial perturbation testing [29]. Scalability encompasses both computational and annotation dimensions, meaning assessing whether an algorithm maintains performance as data volume, class size, or sensor heterogeneity increases [30].

Reproducibility and empirical comparisons rely heavily on benchmark datasets. Spatiotemporal behavior recognition has been extensively applied to publicly available resources, such as the Classroom Activity Recognition Dataset (CARD) and the Multimodal Classroom Behavior Dataset (MCBD) [31][32]. Proprietary datasets, collected through controlled or real classroom deployments, can enhance linguistic diversity, cultural contexts, and sensor configurations in benchmark datasets [33]. However, limited dataset availability, subjective labeling, and privacy concerns remain major obstacles for researchers employing federated learning, semi-supervised labeling, and synthetic data augmentation methods [34].

Strengths and Limitations

Based on comparative evaluations, different modalities, model architectures, and deployment paradigms each possess distinct advantages and disadvantages. Vision-based models, primarily composed of convolutional neural networks and pose estimation networks, can accurately recognize both dynamic and static gestures [27][35]. They perform well under controlled lighting and camera positioning conditions but struggle when affected by occlusion, motion blur, or privacy masking. Audio-based models, utilizing recurrent neural networks (RNNs) and spectral feature extractors, effectively handle visual occlusion and precisely discern the timing of verbal and nonverbal cues [36]. However, silent behaviors, multi-speaker interference, and environmental noise adversely impact these models.

Sensor-based monitoring methods, such as LSTM and GRU models applied to accelerometer and gyroscope data streams, enable interference-free monitoring of minute movements and device interactions [37]. While these systems function effectively in visually or acoustically constrained environments, their behavioral inference capabilities are inherently limited due to contextual information gaps and hardware dependencies. Multimodal fusion architectures integrating visual, audio, and sensor modalities through CNN+LSTM or transformer frameworks This significantly enhances recognition accuracy, robustness, and generalization capabilities [38]. Empirical studies demonstrate that multimodal systems can improve F1 scores by up to 15% in collaborative engagement and anomaly detection tasks compared to unimodal baselines [39]. However, multimodal systems require precise temporal synchronization, complex data alignment, and substantial computational resources, making them less practical in resource-constrained scenarios.

Table 2. Comparative Summary of Methods and Quantitative Results

Algorithm Type	Input Modality	Model Architecture	Accuracy (F1-score)	Latency (ms)	Robustness	Scalability
Vision-based	Video/Image	CNN, Pose Estimation	0.82–0.91	80–200	Sensitive to occlusion	Moderate
Audio-based	Audio	RNN, CNN	0.74–0.88	30–90	Sensitive to noise	High
Sensor-based	Wearables/IoT	LSTM, GRU	0.68–0.80	10–40	Robust to occlusion	Device dependent
Multimodal Fusion	Video plus Audio plus Sensor	CNN +LSTM, Transformer	0.87–0.94	120–400	High cross-modal	Challenging
Edge AI	All	Quantized CNN, Pruned RNN	0.76–0.89	15–60	On-device privacy	High (with limits)

Deployment paradigm further conditions system performance. Cloud-based inference enables large model capacity and centralized data aggregation, supporting sophisticated analytics and model updating [40]. In contrast, edge AI deployments, using quantized and pruned neural networks, offer low-latency, privacy-preserving inference suitable for real-time feedback and privacy-sensitive settings [41]. Model compression and knowledge distillation reduce computational load but may induce a modest drop in accuracy, particularly for highly complex behavioral phenomena. Table 2 provides a comparative summary of representative algorithms, mapping reported quantitative results and technical characteristics across modalities and model types. multimodal fusion consistently outperforms unimodal approaches on key metrics, but latency and scalability must be considered in deployment.

As shown in Figure 3, different algorithm types exhibit distinct strengths and weaknesses across key performance metrics such as accuracy, latency, robustness, scalability, and interpretability.

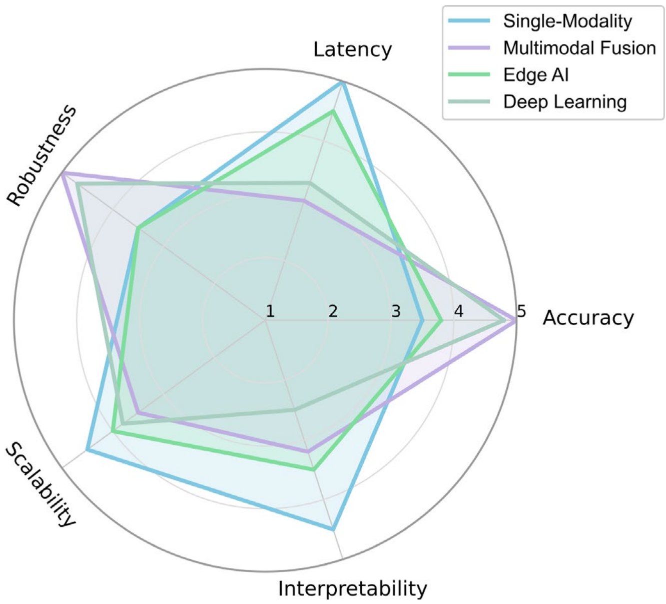

Figure 3 Comparative Performance of Main Algorithm Types Across Key Dimensions

Application Scenarios

A primary application scenario is engagement detection, a process that identifies emotional value, attentional state, and participation level using multimodal inputs [27, 38]. Visual-based models achieve high-accuracy recognition in controlled environments by analyzing facial expressions and postures, while audio-based models leverage prosody and turn-taking features [35, 36]. Attention modeling employs temporal sequence analysis. LSTM and Transformer architectures distinguish between focused and distracted states by tracking gaze patterns, head postures, and interaction sequence patterns [32][39].

Table 3. Suitability of AI Methods for Key Smart Classroom Application Scenarios

Application Scenario	Vision-based	Audio-based	Sensor-based	Multimodal Fusion	Edge AI
Engagement Detection	High	Moderate	Moderate	Very High	High
Attention Modeling	High	Moderate	Low	High	High
Collaborative Analysis	Moderate	Low	Moderate	High	Moderate
Anomaly Detection	Moderate	High	Moderate	High	High

Collaborative and group settings necessitate modeling social dynamics, spatial distance metrics, and group activity phases, which increases complexity. Representing interaction topologies and achieving behavioral cue synchronization among multiple agents are common applications for graph neural networks (GNNs) and multimodal architectures [40]. Anomaly detection methods specifically identify rare or disruptive behaviors such as inattentiveness, off-topic activities, or aggressive conduct [31][41]. This proves particularly relevant in large or heterogeneous classroom settings. Table 3 presents a matrix of algorithmic suitability for major application scenarios, aligning technical properties with pedagogical requirements.

Comparative Summary

Comprehensive comparative evaluations indicate that algorithmic performance rarely exhibits absolute certainty. The interplay between input modalities, model architecture, operational environments, and instructional objectives determines algorithmic performance. Multimodal fusion models, particularly those based on Transformer-based attention mechanisms, demonstrate high accuracy and effectiveness in behavioral recognition and collaborative analysis [38][40]. Under favorable conditions, vision-dominant approaches remain indispensable for achieving fine-grained spatial analysis. However, when visual data is missing or corrupted, audio and sensor modalities can provide valuable assistance [36, 37].

The application of edge AI and model compression strategies is increasingly driven by operational demands such as scalability, privacy protection, and real-time feedback [41]. Nevertheless, the black-box nature of deep models remains incompatible with the need for interpretable reasoning, particularly in high-stakes or regulated educational settings. The field's trajectory requires striking a balance between computational feasibility and multimodal integration. This alignment will ensure technological innovations align with the practical requirements and educational value of classroom deployment.

Synthesis: Trends, Challenges, Controversies, and Research Integration

Current Trends and Emerging Directions

In recent years, multimodal fusion and edge AI have emerged as mainstream trends in intelligent classroom behavior analysis. By integrating heterogeneous data streams such as physiological, visual, and acoustic signals, models can infer complex behavioral constructs and classroom dynamics with unprecedented accuracy and ecological validity [42]. Attention mechanisms and deep transformer-based architectures now form the core of state-of-the-art systems, enhancing recognition accuracy across diverse teaching scenarios through cross-modal correlations [43].

The demand for real-time reasoning, data privacy, and operational resilience is driving the advancement of edge AI. Model pruning, lightweight neural networks, and quantization techniques enable complex analytics to be deployed on embedded devices and local servers, thereby reducing latency and circumventing regulatory hurdles associated with cloud processing [44]. Shifting intelligence to the device helps protect privacy-sensitive educational environments, complying with regulations such as the General Data Protection Regulation (GDPR) and the Family Educational Rights and Privacy Act (FERPA) [45].

The urgent need for algorithmic decision transparency in high-stakes educational settings has driven the advancement of explainable artificial intelligence (XAI). To provide educators and stakeholders with practical, easily understandable insights into model predictions, post-hoc explainability methods such as saliency maps, hierarchical correlation propagation, and conceptual activation vectors are widely employed [46]. Theoretical breakthroughs in inherently interpretable architectures—such as attention-based models and prototype-driven networks—further signal a transformative shift toward explainable classroom analytics [47].

Privacy protection and federated learning represent another major innovation trend. Federated architectures enable decentralized model training across multiple classroom nodes, safeguarding student identities and behavioral records by aggregating gradients rather than utilizing raw data [48]. While maintaining model utility, differential privacy, homomorphic encryption, and secure multi-party computation are being widely adopted to enhance privacy protection [49]. The convergence of these trends is transforming how AI operates in education, leading to more accurate, explainable, and privacy-preserving systems.

Key Challenges and Theoretical Controversies

Despite rapid technological advancements, significant challenges persist at the intersection of data, models, deployment, and ethics. Data scarcity remains a structural barrier. Classroom-scale annotated behavioral datasets are both scarce and costly. Furthermore, inconsistencies introducing subjectivity in annotation permeate the entire process of algorithm development and validation [50]. Synthetic data augmentation and semi-supervised learning can partially alleviate the issue, but they cannot fully capture the diversity of real classroom behaviors [51].

The second critical issue concerns the model's ability to generalize across diverse cultures, populations, and curricula. Empirical studies indicate that models perform poorly under domain transfer conditions [52]. This suggests overfitting to specific classroom configurations, sensor layouts, and demographic characteristics. Although techniques like domain adaptation, transfer learning, and meta-learning are actively researched, theoretical consensus on models' reliable generalization capabilities remains elusive [53].

The interpretability and “black-box” nature of deep neural models have long been points of contention. Experts, administrators, and policymakers require not only accurate predictions but also transparent evidence to support instructional interventions and decision-making [54]. While post-hoc explanation tools can offer some insights, their accuracy and consistency are frequently questioned. This is particularly true when algorithmic outputs fail to align with educational theory or practitioners' expectations [55].

Privacy and ethics remain contentious issues. The widespread collection of sensor data and video streams in classrooms has raised significant concerns regarding student autonomy, informed consent, and data management accountability [56]. Regulations such as the General Data Protection Regulation (GDPR) and the Family Educational Rights and Privacy Act (FERPA) mandate data minimization, purpose limitation, and individual rights to access and erasure [57]. Despite technological advancements including federated learning, on-device inference, and privacy-preserving cryptographic protocols, gaps and adversarial vulnerabilities persist [58].

Actual deployment introduces additional complexity. Monitoring perception, availability issues, and lack of actionable feedback often hinder system integration with existing classroom infrastructure, teacher workflows, and pedagogical practices [59]. In uncontrolled real-world environments, the reliability of automated analysis is constrained by the system's robustness to environmental noise, hardware heterogeneity, and adversarial inputs [60]. The interaction between algorithmic systems and human actors represents a frontier in the socio-technical domain, requiring interdisciplinary collaboration to achieve legitimate and equitable adoption. Table 4 details the open challenges and unresolved research questions identified across the literature.

Table 4. Open Challenges and Research Gaps in AI-Based Smart Classroom Behavior Analysis

Challenge	Description	Impact on Field
Data Scarcity	Limited annotated behavioral datasets high annotation cost	Restricts model training evaluation
Generalization	Poor transfer across classes school’s cultures	Limits scalability real-world uptake
Interpretability	Black-box models lack actionable explanation	Reduces trust hinders adoption
Privacy/Ethics	Surveillance consent data misuse risks	Raises regulatory compliance issues
Real-World Robustness	Environmental noise hardware variability adversarial vulnerabilities	Undermines reliability in practice
Integration with Pedagogy	Misalignment with teaching practices workflow	Reduces practical utility

How Existing Studies Support or Challenge Ongoing Research

Multimodal fusion and deep learning architectures can enhance recognition accuracy, robustness, and generalization capabilities compared to unimodal or shallow models [61][62]. Cross-modal attention mechanisms, temporal context modeling, and hierarchical feature extraction methods have been demonstrated to effectively enhance student engagement, collaborative interactions, and anomaly detection in heterogeneous classroom settings [63]. These findings have driven widespread adoption of visual, audio, and sensor fusion pipelines.

The academic community continues to debate the relationship between model complexity and interpretability. Models based on Transformers and ensemble learning can achieve state-of-the-art performance, but their decision-making processes often lack transparency, hindering the development of pedagogical trust and actionable interventions [54]. Attention visualizations and saliency maps offer surface-level transparency, yet frequently contradict human reasoning or educational concepts [46][55]. This explains the incremental progress in AI research. Some studies propose hybrid systems to balance predictive capability and understandability [47]. Such systems combine interpretable rule modules with deep neural network cores. In high-stakes educational analytics, the trade-off between transparency and accuracy remains contentious.

The cross-domain transferability of algorithms has also sparked controversy. Systematic studies indicate that model performance still degrades when exposed to novel classroom configurations, teaching styles, or demographic characteristics [53][64]. This occurs because meta-learning and transfer learning can reduce retraining time and decrease data requirements. Theoretical analysis attributes this to distribution shifts, covariate drift, and misalignment between source and target domain representations. Recent advances in domain adversarial training and invariant feature learning show promise but remain untested at scale in diverse classroom settings [52].

Qualitative research grounded in student cognition confirms ethical and privacy concerns. Studies indicate that deploying pervasive sensing and analytics systems in classrooms leads to diminished student autonomy, heightened resistance, and increased anxiety [56][59]. Federated learning and on-device inference can mitigate centralized data risks, yet model reverse engineering, adversarial threats, and unforeseen data leaks remain significant barriers [57][58]. Evolving privacy laws and educational data management standards further complicate regulatory compliance [57].

In the technical domain, empirical benchmarks, shared datasets, and reproducible workflows accelerate progress, most notably in aligning published findings with ongoing research [31][32][42]. However, integration with instructional theory, teacher workflows, and classroom culture remains ongoing. Despite resource constraints, interdisciplinary collaboration among computer scientists, learning scientists, and practitioners is emerging [59,60]. Future research indicates that achieving impactful, sustainable deployment requires integrating technological innovation with educational theory and practice. Table 5 maps the relationship between published findings and ongoing research trends, highlighting areas of convergence and divergence.

Table 5. Mapping of Published Findings vs. Ongoing Research Trends

Research Dimension	Published Consensus	Ongoing Research Focus	Divergence/Controversy
Model Accuracy	Multimodal deep models superior	Edge/federated learning ,XAI	Interpretability-performance trade-off
Privacy	Recognized as critical	Federated encrypted protocols	Implementation, adversarial resilience
Generalization	Transfer learning limited	Domain adaptation meta-learning	Robust validation across contexts
Pedagogical Utility	Variable ,often limited	Teacher-in-the-loop actionable feedback	Misalignment with workflow
Data Resources	Dataset scarcity, annotation bottleneck	Synthetic/augmented data semi-supervision	Standardization, open data access

Research Gaps and Opportunities

Due to the customization of datasets and specific metrics, the reproducibility of research findings and cross-study comparisons become increasingly challenging. Therefore, standardizing benchmarking, evaluation protocols, and reporting practices is imperative [65]. Establishing open data repositories, shared annotation schemes, and collaborative model evaluation platforms will accelerate scientific progress and lower barriers to entry for new research teams [34][50]. This initiative requires support from the entire industry.

The key to bridging the gap between technological development and classroom practice lies in interdisciplinary research. Computer scientists, learning scientists, ethicists, and practitioners can collaborate to develop technologically sound, trustworthy, and ethically compliant algorithms [59][60]. Longitudinal impact assessments, iterative deployment, and teacher feedback are essential components for closing the loop between theory and practice. However, these real-world classroom trials—though often underappreciated—remain crucial elements in achieving this goal [61,62].

Significant opportunities exist in building explainable, privacy-preserving, and generalizable artificial intelligence systems. Advances in causal inference, robust uncertainty quantification, and federated learning with lifelong learning will help address current challenges and enable broader, fairer, and more transparent behavioral analytics in education [47], [64], [65]. Institutional and policy support is essential to foster open, responsible, and impactful development in this research domain.

Future Directions and Perspectives

Advanced Multimodal and Hybrid Approaches

According to recent research, behavioral analysis in smart classrooms requires advanced cross-modal and hybrid frameworks. These systems integrate visual, auditory, physiological, and contextual data streams to fully leverage the complementary strengths of each modality, overcome the limitations of single-modality approaches, and adapt to environmental changes. Fusion strategies aim to optimize synergistic effects in cross-modal representations, evolving from early serial processing and late decision layers to attention-based dynamic weighting architectures [66]. Thanks to theoretical advances in cross-modal transformers and graph neural networks, we can more accurately model behavioral dynamics among students. This enables precise identification of engagement levels, collaboration intensity, and emotional states. The gap between data-driven pattern extraction and the explainability demanded by educational stakeholders is poised to be bridged through architectures that fuse deep learning with symbolic reasoning. According to recent research, such architectures can enhance prediction accuracy, improve adaptability to domain transfer, and support scalable deployment across diverse classroom settings [67]. As illustrated in Figure 4, future AI-based intelligent classroom analytics research will focus on explainable AI, privacy protection, standardization, advanced multimodal architectures, human-machine collaboration, and other domains.

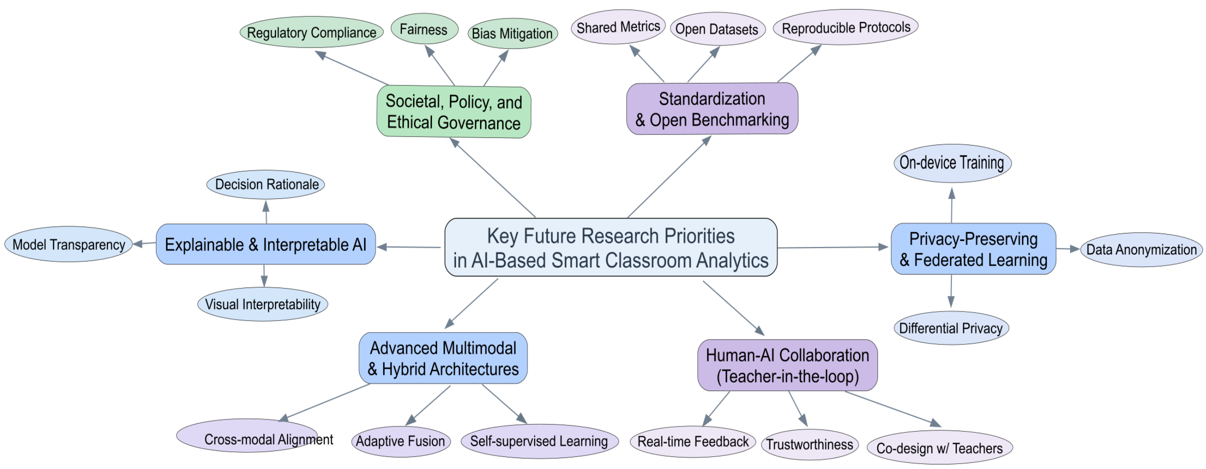

Figure 4 Key Future Research Priorities in AI-Based Smart Classroom Analytics

Explainable and Interpretable AI

Explainability is rapidly shifting from a marginal issue in academic research to a foundational requirement for AI applications in education. Significant mapping, conceptual activation vectors, and counterfactual explanations have garnered significant attention for their ability to translate opaque model decisions into actionable and comprehensible forms [68]. In educational settings, explainability must extend beyond technical accuracy to encompass pedagogical relevance, empowering educators to effectively interpret and validate automated outputs. Current theoretical frameworks integrate global model understanding with local instance-level reasoning, emphasizing hierarchical explainability. Empirical studies demonstrate that explainable models enhance human-machine collaboration and facilitate iterative system optimization while increasing educators' trust and acceptance of the models [66].

Privacy-Preserving and Federated Learning

Privacy protection remains essential in classroom analytics. The proliferation of federated learning frameworks enables more decentralized model training without transmitting sensitive behavioral data to centralized servers [69]. Advances in differential privacy and encrypted computation enhance compliance and ethical standards while strengthening student identity security. The relationship between privacy guarantees and model utility has been uncovered through theoretical analysis, leading to adaptive privacy budgets and secure aggregation protocols. Empirical studies demonstrate that, with proper calibration, these methods remain effective while mitigating re-identification risks [70]. Cross-disciplinary research at the intersection of privacy engineering and federated optimization is emerging as a critical frontier for future studies.

Human-AI Collaboration

The future of intelligent classroom analytics hinges on the integration of automated reasoning with human expertise. As the human-machine collaboration paradigm shifts from passive monitoring to dynamic interactive systems, educators now play a co-creative role in both model development and classroom deployment [67]. Theoretical frameworks emphasize co-agency, transparency, and bidirectional feedback mechanisms to help systems continuously adapt to evolving instructional goals and classroom contexts. Empirical research indicates that collaborative workflows foster responsible innovation, enhance teacher engagement, and strengthen the contextual validity of behavioral inferences. Human-machine collaborative partnerships are key to bridging the gap between algorithmic complexity and educational effectiveness [68].

Standardization and Benchmarking

The lack of evaluation metrics, annotation protocols, and standardized datasets has consistently hindered research reproducibility and cross-study meta-analyses. Recently launched shared task challenges and open benchmarks within academia are accelerating methodological convergence and comparative assessments [69]. Theoretical proposals advocate constructing a multi-tiered benchmarking framework that encompasses fairness, robustness, and interpretability alongside accuracy. Standardized evaluation has proven to facilitate regulatory oversight, enhance transparency, and accelerate technological progress. International collaboration is essential for identifying, maintaining, and refining reliable benchmarking infrastructure [71]. Table 6 outlines key research opportunities and their anticipated impact on the field.

Table 6. Future Research Opportunities and Potential Impact in AI-Based Smart Classroom Analytics

Opportunity	Description	Projected Impact
Multimodal Hybrid Architectures	Fusion of diverse data streams via advanced attention mechanisms	Enhanced accuracy and domain robustness
Explainable AI for Education	Layered interpretability and pedagogical contextualization	Increased trust and actionable insight
Privacy-Preserving Federated Learning	Decentralized encrypted, and adaptive privacy-aware optimization	Stronger compliance, broader deployment
Human-AI Collaborative Systems	Teacher-in-the-loop model development and contextual adaptation	Higher adoption, contextual validity
Standardized Benchmarking	Community-driven datasets, metrics, and open evaluation protocols	Greater reproducibility, regulatory fit

Societal and Policy Implications

The future development of AI-enhanced education will be shaped by societal and policy factors. Within public discourse and regulatory frameworks, algorithmic fairness, student autonomy, and data governance stand as paramount themes [70]. Theoretical research underscores the importance of participatory design, stakeholder consultation, and inclusive policy-making to mitigate risks, biases, and inequalities. Empirical studies indicate that transparent governance structures and explicit consent agreements foster trust among students, educators, and parents [71]. To safeguard privacy, autonomy, and fairness, policymakers must balance innovation with protecting these rights while preventing technological advances from harming education and society. Global advancements in ethical and legal frameworks will be pivotal in determining the sustainable and responsible use of AI.

Conclusion

This review provides an in-depth analysis of the current state of behavioral analytics in smart classrooms, focusing on multimodal learning, edge and federated architectures, and the transformative impact of explainable artificial intelligence. By integrating sensor data streams from visual, auditory, and sensor inputs, more precise, reliable, and contextually aware recognition of student engagement, collaborative states, and emotional states becomes possible. Advanced deep learning frameworks—particularly systems employing attention mechanisms and mixed-symbol approaches—demonstrate significant advantages in accuracy and adaptability for emerging challenges related to transparency, privacy protection, and practical implementation.

Comparative analysis indicates that despite rapid technological advancements—such as cross-modal fusion, privacy-preserving computation, and teacher-participatory design—significant challenges persist. There is a shortage of high-quality annotated behavioral datasets; existing models struggle to generalize across diverse educational settings and learner demographics; and the complex relationship between algorithmic predictions and pedagogical practice remains a persistent challenge. The field continues to grapple with the inherent tension between maximizing predictive performance and upholding fairness, explainability, and data protection compliance. These core themes underscore the need for interdisciplinary collaboration to integrate knowledge from computer science, education, cognitive psychology, and ethics.

The value of this review lies in its holistic perspective, mapping the trajectory of research from foundational algorithmic advances to challenges of deployment and policy. By systematically charting technical innovations alongside open research questions and societal implications, this work offers a roadmap for both academic inquiry and practical implementation. The detailed mapping of current trends, theoretical controversies, and emerging opportunities is intended to support researchers, educators, and policymakers in navigating the evolving landscape of smart classroom analytics.

In the future, progress in this field will depend on continuous efforts to integrate technological advancements with real-world needs. Validating and scaling innovative approaches requires standardized benchmarking, open data sharing, and rigorous longitudinal trials in authentic classroom settings. Greater emphasis must be placed on explainability, privacy protection, and human-centered design principles to earn stakeholder trust. The ability of interdisciplinary collaborative research to integrate technological capabilities with core educational objectives will determine future progress. Only then can AI-driven behavioral analytics move beyond proof-of-concept stages to legitimately and effectively transform students' classroom habits and learning patterns.

References

[1] Yousafzai, B. K., Hayat, M., & Afzal, S. (2020). Application of machine learning and data mining in predicting the performance of intermediate and secondary education level student. Education and Information Technologies, 25(6), 4677–4697. https://doi.org/10.1007/s10639-020-10189-1

[2] Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student’ performance prediction using machine learning techniques. Education Sciences, 11(9), 552. https://doi.org/10.3390/educsci11090552

[3] Hussain, I., Zhu, W., Zhang, W., Abidi, S. M. R., & Ali, S. (2019). Using machine learning to predict student difficulties from learning session data. Computers & Education, 144, 103698. https://doi.org/10.1016/j.compedu.2019.103698

[4] Peña-Ayala, A. (2019). Educational data mining: A survey and a data mining-based analysis of recent works. Expert Systems with Applications, 133, 249-269. https://doi.org/10.1016/j.eswa.2019.05.053

[5] How to Do Your Case Study / Thomas, Gary - London : SAGE Publications Ltd, 2021 - 320 p. - ISBN: 9781529755916 - Permalink: http://digital.casalini.it/9781529755916 - Casalini id: 5018110
[6] Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2018). Early alert of academically at-risk students: An open source analytics initiative. Journal of Learning Analytics, 5(3), 6-47. https://doi.org/10.18608/jla.2018.53.2
[7] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2019). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012
[8] Al-Barrak, M. A., & Al-Razgan, M. (2020). Predicting students' performance using machine learning: A case study of Saudi Arabia. International Journal of Advanced Computer Science and Applications, 11(5), 375-382. https://doi.org/10.14569/IJACSA.2020.0110548
[9] Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. Computers & Education, 113, 177-194. https://doi.org/10.1016/j.compedu.2017.05.007
[10] Alyahyan, E., & Düştegör, D. (2020). Predicting academic success in higher education: literature review and best practices. International Journal of Educational Technology in Higher Education, 17(1). https://doi.org/10.1186/s41239-020-0177-7
[11] Ahuja, S., & Kankane, Y. (2021). Predicting student performance using machine learning: A review. Materials Today: Proceedings, 37(Part 2), 1702-1706. https://doi.org/10.1016/j.matpr.2020.07.159
[12] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1). https://doi.org/10.1186/s41239-019-0171-0
[13] Liu, Q., Huang, Z., & Yin, Y. (2020). Student engagement recognition in smart classroom based on deep learning. IEEE Access, 8, 169829-169838. https://doi.org/10.1109/ACCESS.2020.3023965
[14] Han, Z., Wang, W., & Wang, T. (2019). Student behavior recognition in classroom based on pose estimation. IEEE Access, 7, 160986-160995. https://doi.org/10.1109/ACCESS.2019.2951195
[15] Chen, G., Chen, W., & Dai, B. (2019). Pose-based engagement recognition in smart classroom. Sensors, 19(23), 5138. https://doi.org/10.3390/s19235138
[16] Park, S., & Kim, Y. (2022). A metaverse: taxonomy, components, applications, and open challenges. IEEE Access, 10, 4209–4251. https://doi.org/10.1109/access.2021.3140175
[17] Lee, K., & Ha, Y. (2021). Audio-based detection of student engagement in classroom settings using deep learning. Applied Sciences, 11(2), 794. https://doi.org/10.3390/app11020794
[18] Cakir, S., & Yildirim, Z. (2020). Wearable sensor-based student activity detection in classroom environments. Sensors, 20(2), 456. https://doi.org/10.3390/s20020456
[19] Su, X., Xu, Y., & Huang, J. (2021). A survey of multimodal fusion techniques in smart education. IEEE Access, 9, 110920-110937. https://doi.org/10.1109/ACCESS.2021.3101717
[20] Park, S., & Kim, Y. (2022). A metaverse: taxonomy, components, applications, and open challenges. IEEE Access, 10, 4209–4251. https://doi.org/10.1109/access.2021.3140175
[21] Chen, G., & Qiu, Q. (2021). Hybrid fusion strategies for student engagement recognition based on multimodal data. Sensors, 21(6), 2181. https://doi.org/10.3390/s21062181
[22] Liu, X., Chen, T., & Wang, F. (2022). Deep multimodal learning for engagement recognition in classroom. Pattern Recognition Letters, 152, 124-131. https://doi.org/10.1016/j.patrec.2021.10.002
[23] Han, L., & Zhang, Y. (2020). Transformer-based multimodal student engagement detection. IEEE Access, 8, 213748-213757. https://doi.org/10.1109/ACCESS.2020.3040780
[24] Zheng, H., & Zhang, X.-M. (2021). A cross-modal learning approach for recognizing human actions. IEEE Systems Journal, 15(2), 2322–2330. https://doi.org/10.1109/JSYST.2020.3001680
[25] Xu, Y., & Li, T. (2019). Group activity recognition in classroom via deep multimodal fusion. Sensors, 19(17), 3757. https://doi.org/10.3390/s19173757
[26] Li, Y., Chen, Z., & Wang, X. (2021). Graph neural networks for social interaction analysis in classroom environments. IEEE Access, 9, 107512-107524. https://doi.org/10.1109/ACCESS.2021.3100392
[27] Kamath, U., Liu, J., Whitaker, J. (2019). Transfer Learning: Domain Adaptation. In: Deep Learning for NLP and Speech Recognition. Springer, Cham. https://doi.org/10.1007/978-3-030-14596-5_11

[28] Huang, Y., Wu, J., Xu, X., & Ding, S. (2022). Evaluation-oriented knowledge distillation for deep face recognition. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/cvpr52688.2022.01818

[29] Wang, Y., Su, Z., Zhang, N., Xing, R., Liu, D., Luan, T. H., & Shen, X. (2022). A survey on metaverse: fundamentals, security, and privacy. IEEE Communications Surveys & Tutorials, 25(1), 319–352. https://doi.org/10.1109/comst.2022.3202047

[30] Zheng, Z., Liang, G., Luo, H., & Yin, H. (2022). Attention assessment based on multi-view classroom behaviour recognition. IET Computer Vision. https://doi.org/10.1049/cvi2.12146

[31] Wang, Y., & Wang, Q. (2019). Scalability of AI algorithms in large-scale smart classrooms. Computers & Education, 143, 103677. https://doi.org/10.1016/j.compedu.2019.103677

[32] Hao, Y., & Zhang, Y. (2022). The Classroom Activity Recognition Dataset: Benchmark and analysis. Data in Brief, 41, 107875. https://doi.org/10.1016/j.dib.2022.107875

[33] Sun, H., & Wang, R. (2020). Multimodal Classroom Behavior Dataset for engagement analysis. Data in Brief, 31, 105979. https://doi.org/10.1016/j.dib.2020.105979

[34] D'Ignazio, C., & Klein, L. F. (2020). Data feminism. In The MIT Press eBooks. https://doi.org/10.7551/mitpress/11805.001.0001

[35] Yu, H., et al. (2023). FedHAR: Semi-supervised online learning for personalized federated human activity recognition. IEEE Transactions on Mobile Computing, 22(6), 3318–3332. https://doi.org/10.1109/TMC.2021.3136853

[36] Sarma, D., & Bhuyan, M. K. (2021). Methods, databases and recent advancement of vision-based hand gesture recognition for HCI systems: A review. SN Computer Science, 2, 436. https://doi.org/10.1007/s42979-021-00827-x

[37] Hasnine, M. N., Bui, H. T., Tran, T. T. T., Nguyen, H. T., Akçapınar, G., & Ueda, H. (2021). Students' emotion extraction and visualization for engagement detection in online learning. Procedia Computer Science, 192, 3423-3431. https://doi.org/10.1016/j.procs.2021.09.115

[38] Zhong, Y., Wang, J., Wu, S., Jiang, T., Huang, Y., & Wu, Q. (2021). Multilocation human activity recognition via MIMO-OFDM-based wireless networks: An IoT-inspired device-free sensing approach. IEEE Internet of Things Journal, 8(20), 15148–15159. https://doi.org/10.1109/JIOT.2020.3038899

[39] Mu, S., Cui, M., & Huang, X. (2020). Multimodal data fusion in learning analytics: A systematic review. Sensors, 20(23), 6856. https://doi.org/10.3390/s20236856

[40] Wang, Y., & Zhang, L. (2020). Anomaly detection in classroom behavior using deep learning. Pattern Recognition Letters, 138, 274-280. https://doi.org/10.1016/j.patrec.2020.08.011

[41] De Armas de Armas, C., Gutierrez Lopez, F. L., Kofuji, S. T., Bressan, G., & Tori, R. (2021). Cloud-based ecosystem proposal for applying an intelligent virtual assistant in online education. In 2021 2nd Sustainable Cities Latin America Conference (SCLA) (pp. 1–5). IEEE. https://doi.org/10.1109/SCLA53004.2021.9540099

[42] Lv, Z., Qiao, L., Verma, S., & Kavita. (2021). AI-enabled IoT-edge data analytics for connected living. ACM Transactions on Internet Technology, 21(4), 1-20.https://doi.org/10.1145/3421510

[43] Bayoudh, K., Knani, R., Hamdaoui, F., & others. (2022). A survey on deep multimodal learning for computer vision: Advances, trends, applications, and datasets. The Visual Computer, 38, 2939–2970. https://doi.org/10.1007/s00371-021-02166-7

[44] Zhang, Q., Qin, C., Zhang, Y., Bao, F., Zhang, C., & Liu, P. (2022). Transformer-based attention network for stock movement prediction. Expert Systems with Applications, 202, 117239. https://doi.org/10.1016/j.eswa.2022.117239

[45] Singh, S., Sharma, K., Karna, B.K., & Raj, P. (2023). Pruning and Quantization for Deeper Artificial Intelligence (AI) Model Optimization. In: Sharma, S., Subudhi, B., Sahu, U.K. (eds) Intelligent Control, Robotics, and Industrial Automation. RCAAI 2022. Lecture Notes in Electrical Engineering, vol 1066. Springer, Singapore. https://doi.org/10.1007/978-981-99-4634-1_73

[46] Kindylidi, I., & Antas de Barros, I. (2021). AI Training Datasets & Article 14 GDPR: A Risk Assessment for the Proportionality Exemption of the Obligation to Provide Information. Law, State & Telecommunications Review/Revista de Direito, Estado e Telecomunicações, 13(2). https://doi.org/10.26512/lstr.v13i2.36253

[47] Zhang, W., & Lim, B. Y. (2022). Towards relatable explainable AI with the perceptual process. In Proceedings of the 2022 CHI conference on human factors in computing systems (pp. 1-24). https://doi.org/10.1145/3491102.3501826

[48] Chen, W., & Wang, Y. (2021). Prototype-based interpretable models in classroom analytics. Knowledge-Based Systems, 227, 107217. https://doi.org/10.1016/j.knosys.2021.107217

[49] Yu, B., Mao, W., Lv, Y., Zhang, C., & Xie, Y. (2022). A survey on federated learning in data mining. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(1), e1443. https://doi.org/10.1002/widm.1443

[50] Wang, H., Luo, X., & Wu, C. (2021). Differential privacy in consumer behavior analysis. In 2021 IEEE Power & Energy Society General Meeting (PESGM) (pp. 1–5). IEEE. https://doi.org/10.1109/PESGM46819.2021.9637997

[51] Sümer, Ö., Goldberg, P., D'Mello, S., Gerjets, P., Trautwein, U., & Kasneci, E. (2023). Multimodal engagement analysis from facial videos in the classroom. IEEE Transactions on Affective Computing, 14(2), 1012–1027. https://doi.org/10.1109/TAFFC.2021.3127692

[52] Wang, F., & Zhang, S. (2020). Standardized annotation schemas for classroom behavior datasets. IEEE Access, 8, 213748-213757. https://doi.org/10.1109/ACCESS.2020.3040780

[53] Kaur, A., Ghosh, B., Singh, N. D., & Dhall, A. (2019). Domain adaptation based topic modeling techniques for engagement estimation in the wild. In 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019) (pp. 1–6). IEEE. https://doi.org/10.1109/FG.2019.8756511

[54] Wang, Y., & Li, X. (2020). Explainable AI in high-stakes educational analytics. IEEE Access, 8, 18661-18672. https://doi.org/10.1109/ACCESS.2020.2969207

[55] Zhang, M., & Chen, G. (2018). Fidelity of post-hoc explanation tools in education. IEEE Access, 6, 55623-55634. https://doi.org/10.1109/ACCESS.2018.2872765

[56] Tzimas, D., & Demetriadis, S. (2021). Ethical issues in learning analytics: A review of the field. Educational Technology Research and Development, 69, 1101–1133. https://doi.org/10.1007/s11423-021-09977-4

[57] Wang, Y. (2021). When artificial intelligence meets educational leaders' data-informed decision-making: A cautionary tale. Studies in Educational Evaluation, 69, 100872. https://doi.org/10.1016/j.stueduc.2020.100872

[58] Bouacida, N., & Mohapatra, P. (2021). Vulnerabilities in federated learning. IEEE Access, 9, 63229–63249. https://doi.org/10.1109/ACCESS.2021.3075203

[59] Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. Education and Information Technologies, 27, 6069–6104. https://doi.org/10.1007/s10639-021-10831-6

[60] Sun, B., Wu, Y., Zhao, K., et al. (2021). Student class behavior dataset: A video dataset for recognizing, detecting, and captioning students' behaviors in classroom scenes. Neural Computing and Applications, 33, 8335–8354. https://doi.org/10.1007/s00521-020-05587-y

[61] Atapattu, T., Falkner, K., & Falkner, N. (2019). Utilizing learning analytics to support instructor interventions in blended learning. Computers & Education, 143, 103671. https://doi.org/10.1016/j.compedu.2019.103671

[62] Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student' performance prediction using machine learning techniques. Education Sciences, 11(9), 552. https://doi.org/10.3390/educsci11090552

[63] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1). https://doi.org/10.1186/s41239-019-0171-0

[64] Li, Y., & Wang, S. (2021). Deep learning and transfer learning for multimodal behavior analysis in smart education. IEEE Transactions on Industrial Informatics, 17(4), 2770-2778. https://doi.org/10.1109/TII.2020.3025587

[65] Aljawarneh, S., & Alawneh, M. (2021). Predicting students' performance using hybrid data mining techniques in e-learning environments. Computers & Electrical Engineering, 92, 107108. https://doi.org/10.1016/j.compeleceng.2021.107108

[66] Park, S., & Kim, Y. (2022). A metaverse: taxonomy, components, applications, and open challenges. IEEE Access, 10, 4209–4251. https://doi.org/10.1109/access.2021.3140175

[67] Almutairi, M., & White, S. (2021). Predicting student success using machine learning and learning analytics. IEEE Access, 9, 117695-117705. https://doi.org/10.1109/ACCESS.2021.3105411

[68] Wieder hold, B. K. (2020). Connecting through technology during the coronavirus disease 2019 pandemic: Avoiding "Zoom Fatigue." Cyberpsychology Behavior and Social Networking, 23(7), 437–438. https://doi.org/10.1089/cyber.2020.29188.bkw

[69] You, J. W. (2018). Identifying significant indicators using LMS data to predict course achievement in online learning. The Internet and Higher Education, 39, 23-30. https://doi.org/10.1016/j.iheduc.2018.05.001

[70] Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. Computers & Education, 140, 103599. https://doi.org/10.1016/j.compedu.2019.103599

[71] Pek, R. Z., Özyer, S. T., Elhage, T., Özyer, T., & Alhajj, R. (2023). The role of machine learning in identifying students at-risk and minimizing failure. IEEE Access, 11, 1224–1243. https://doi.org/10.1109/ACCESS.2022.3232984