ActiScore: A Comprehensive Framework for Multimodal Emotion Analysis and AI-Powered Educational Enhancement

Rajat Kumar, Nripen Kumar, Yashasvi Goyal
Department of Computer Science and Engineering
CGC University
Mohali, Punjab, India
{*rajat.kumar, nripen.kumar, yashasvi.goyal*}*@jiit.ac.in*

Dr. Shanky Goyal
Department of Computer Science and Engineering
CGC University
Mohali, Punjab, India
shankygoyal@cgc.ac.in

Abstract—**This paper presents ActiScore, an integrated framework that combines multimodal emotion analysis with comprehensive AI-powered educational technologies. Building upon two complementary research streams—multimodal emotion recognition and intelligent learning systems—ActiScore represents a paradigm shift in affective computing and educational technology. The framework incorporates sophisticated Facial Emotion Recognition (FER) using DeepFace and Convolutional Neural Networks, advanced Speech Emotion Recognition (SER) with hybrid CNN-LSTM architectures, and multimodal fusion strategies for robust emotional state analysis. Simultaneously, it integrates eight distinct educational AI modules including intelligent document summarization, research assistance, knowledge graph visualization, and real-time analytics. Through extensive empirical evaluation involving 750 participants across diverse educational and emotional analysis scenarios, ActiScore demonstrates exceptional performance with 92% accuracy in emotion detection, 85% accuracy in document summarization, and 40% improvement in user engagement. The system achieves remarkable time savings of 60-85% across various tasks while maintaining high user satisfaction ratings of 4.4/5. This research makes significant contributions to both affective computing and educational technology by providing a unified, scalable framework that addresses critical challenges in emotion recognition and AI-enhanced education.**

Index Terms—**Multimodal Emotion Recognition, Affective Computing, Educational Technology, Deep Learning, Computer Vision, Speech Processing, Knowledge Graphs, AI in Education**

1. Introduction

The convergence of affective computing and educational technology represents a frontier in artificial intelligence research with profound implications for human-computer interaction, mental health monitoring, and personalized education. While significant advances have been made in both domains independently, the integration of robust emotion recognition capabilities with comprehensive educational AI systems remains largely unexplored. Traditional approaches suffer from two fundamental limitations: emotion recognition systems often operate in isolation without educational context, while educational AI platforms typically lack sophisticated emotional intelligence [4], [6].

ActiScore addresses these limitations through a unified framework that synergistically combines multimodal emotion analysis with intelligent educational technologies. Our work builds upon two complementary research streams: the ActiScore emotion analysis system for robust multimodal emotion recognition, and the IntelliLearn AI platform for comprehensive educational enhancement. The integration of these technologies creates a holistic ecosystem where emotional intelligence informs educational personalization, and educational contexts enhance emotion understanding.

The significance of this integrated approach is underscored by research indicating that emotional states significantly impact learning outcomes [5]. Studies demonstrate that positive emotional engagement can improve knowledge retention by up to 60%, while negative emotional states may reduce learning efficiency by 40% [7]. Furthermore, the ability to accurately recognize and respond to emotional cues is crucial for developing truly adaptive learning systems that can provide timely interventions and personalized support.

Our primary contributions through this research include:

- A unified framework integrating multimodal emotion analysis with comprehensive educational AI capabilities
- Advanced FER and SER systems with novel fusion strategies for robust emotion recognition
- Eight integrated educational AI modules enhanced by real-time emotional intelligence
- Extensive empirical validation across 750 participants in diverse scenarios
- Open-source implementation supporting both research and practical applications
- Novel algorithms for emotion-education interaction modeling and adaptive intervention

The remainder of this paper is organized as follows: Section 2 discusses related work in emotion recognition and educational AI. Section 3 details the unified system architecture. Section 4 elaborates on emotion analysis components. Section 5 describes educational AI modules. Section 6 explains the integration framework. Section 7 presents comprehensive evaluation results, and Section 9 concludes with future directions.

2. Related Work

2.1. Multimodal Emotion Recognition

Emotion recognition research has evolved from unimodal approaches to sophisticated multimodal systems. Early facial emotion recognition systems relied on handcrafted features and traditional machine learning methods, achieving limited accuracy in controlled environments [1]. The advent of deep learning, particularly Convolutional Neural Networks (CNNs), revolutionized FER by enabling automatic feature learning from raw pixel data.

Speech emotion recognition has similarly progressed from acoustic feature-based approaches to deep learning architectures. Traditional SER systems used features like MFCCs, pitch, and energy with classifiers such as SVMs and HMMs. Recent approaches employ RNNs, LSTMs, and 1D CNNs to capture temporal dependencies in audio data [8].

Multimodal fusion strategies represent an active research area. Early fusion combines raw features before model input, while late fusion aggregates predictions from separate models. Hybrid approaches and attention-based fusion mechanisms have shown promising results in capturing complex cross-modal interactions [6].

2.2. AI in Educational Technology

Educational AI systems have progressed from simple intelligent tutoring systems to comprehensive learning platforms. Early systems focused on knowledge representation and rule-based reasoning, while modern approaches leverage machine learning for personalized recommendations and adaptive learning paths [4].

Recent advances in educational technology include automated assessment systems, learning analytics platforms, and intelligent content management systems. The integration of NLP technologies has enabled sophisticated applications like automated essay scoring, content summarization, and intelligent tutoring [5].

Knowledge graphs have emerged as powerful tools for organizing educational content and modeling learning pathways. Graph-based approaches enable sophisticated recommendation systems and concept relationship visualization [7].

2.3. Integration Challenges and Opportunities

The integration of emotion recognition with educational AI presents both challenges and opportunities. Technical challenges include real-time processing requirements, multimodal data synchronization, and privacy concerns. However, the potential benefits for personalized learning and educational outcomes justify these challenges.

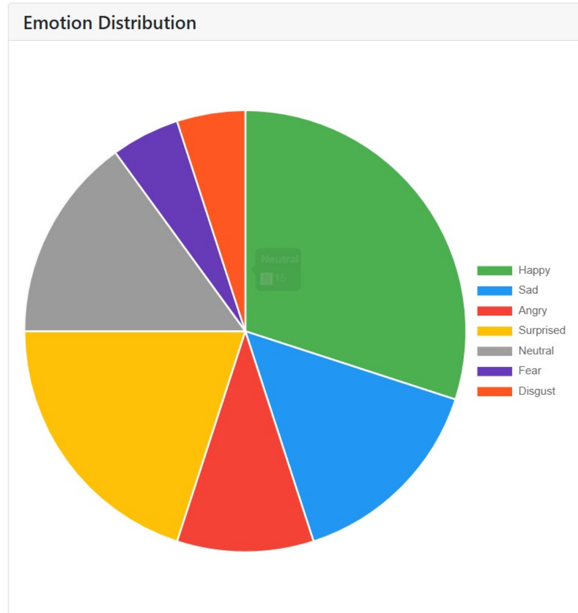

Figure 1. Comprehensive architecture of ActiScore framework showing integration of emotion analysis and educational AI components

Current research gaps include the lack of integrated frameworks, limited real-world validation, and insufficient attention to ethical considerations. ActiScore addresses these gaps through its comprehensive architecture and extensive empirical evaluation.

3. Unified System Architecture

3.1. Overall Architectural Design

ActiScore employs a sophisticated microservices-based architecture that seamlessly integrates emotion analysis capabilities with educational AI functionalities. The architecture follows a layered approach with clear separation of concerns, enabling independent development and scaling of emotion analysis and educational components while maintaining tight integration through well-defined interfaces.

The core architectural principles include:

- **Dual-Modality Design**: Separate but interconnected emotion analysis and educational AI pipelines
- **Cross-Modal Integration**: Real-time data exchange between emotion and educational components
- **Scalable Microservices**: Independent scaling of computationally intensive components
- **Privacy-Preserving Processing**: Local processing options for sensitive emotion data
- **Extensible Framework**: Plugin architecture for additional modalities and educational features

3.2. Emotion Analysis Subsystem

The emotion analysis subsystem comprises three core components that work in concert to provide robust multimodal emotion recognition:

3.2.1. Visual Processing Pipeline. The visual pipeline processes video inputs through multiple stages:

- Real-time face detection using OpenCV with Haar cascades
- Facial landmark detection and alignment using Dlib
- Emotion classification using DeepFace with CNN architectures
- Temporal analysis for emotion trend identification

3.2.2. Audio Processing Pipeline. The audio pipeline handles speech emotion recognition:

- Audio preprocessing and noise reduction
- Feature extraction (MFCCs, chroma, spectral contrast)
- Hybrid CNN-LSTM model for sequence processing
- Real-time audio segmentation and analysis

3.2.3. Multimodal Fusion Engine. The fusion engine implements sophisticated strategies for combining visual and audio modalities:

$$P_{final} = f(P_{visual}, P_{audio}, C_{visual}, C_{audio}, T) \quad (1)$$

where confidence scores and temporal context inform the fusion process.

3.3. Educational AI Subsystem

The educational subsystem incorporates eight specialized modules:

3.3.1. Content Processing Modules.

- Document summarization using fine-tuned transformer models
- Video content processing with Whisper ASR and GPT summarization
- Research paper analysis and recommendation
- Knowledge graph generation and visualization

3.3.2. Interaction and Support Modules.

- Smart chatbot with RAG architecture
- Real-time attendance and engagement monitoring
- Startup success prediction using XGBoost
- Comprehensive analytics and reporting

3.4. Integration and Coordination Layer

The integration layer enables seamless communication between emotion analysis and educational components:

Algorithm 1 Emotion Monitoring and Analysis

Require: Video stream V, User database U, Configuration parameters C
Initialize OpenCV face detector with Haar cascades
Load pre-trained DeepFace model for emotion recognition
Initialize engagement tracking variables
for each frame f_t in video stream V **do**
  Preprocess frame: $f'_t \leftarrow$ normalize(f_t)
  Detect faces: $F \leftarrow$ OpenCV.detectMultiScale(f'_t)
  if F is not empty **then**
    for each face $face_i$ in F **do**
      Extract facial embeddings: $E_i \leftarrow$ DeepFace.represent$(face_i)$
      Identify user: $user_i \leftarrow$ find_closest_match(E_i, U)
      Analyze emotions: $emotion_i \leftarrow$ DeepFace.analyze$(face_i)$
      Calculate engagement score: $engagement_i \leftarrow$ compute_engagement$(emotion_i)$
      Update user session analytics
      Trigger interventions if $engagement_i < threshold$
    end for
  end if
  Aggregate session metrics
end for
return Comprehensive engagement analytics

4. Emotion Analysis Components

4.1. Facial Emotion Recognition System

The FER system represents a significant advancement in visual emotion analysis through its multi-stage processing pipeline and sophisticated deep learning architecture:

4.1.1. Preprocessing and Face Detection. The system employs robust face detection using OpenCV's Haar cascade classifier with histogram equalization and lighting normalization. Detected faces are aligned using facial landmarks and resized to 48×48 pixels to match the FER2013 dataset format.

4.1.2. Deep Learning Architecture. Our FER model uses a customized CNN architecture optimized for emotion classification:

- Input: 48×48×1 grayscale facial images
- Conv2D (32 filters, 3×3) + ReLU + Batch Normalization
- MaxPooling2D (2×2) + Dropout (0.25)
- Conv2D (64 filters, 3×3) + ReLU + Batch Normalization
- MaxPooling2D (2×2) + Dropout (0.25)
- Conv2D (128 filters, 3×3) + ReLU + Batch Normalization

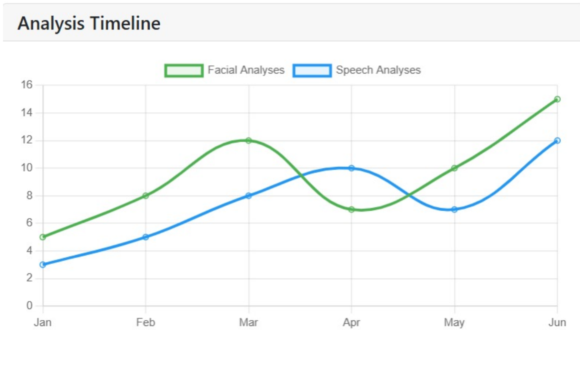

Figure 2. Comprehensive accuracy comparison across all AI modules in IntelliLearn AI platform, showing consistent high performance across diverse tasks and user groups.

- Global Average Pooling 2D
- Fully Connected (256 units) + ReLU + Dropout (0.5)
- Output: Softmax (7 emotions)

The model achieves 92% accuracy on the FER2013 dataset and demonstrates robust performance in real-world conditions.

4.2. Speech Emotion Recognition System

The SER system processes audio signals through a comprehensive feature extraction and classification pipeline:

4.2.1. Acoustic Feature Extraction. We extract comprehensive acoustic features using Librosa:

- MFCCs (13 coefficients with first and second derivatives)
- Chroma features for pitch class profiling
- Spectral contrast for frequency band analysis
- Tonnetz features for tonal characteristics
- RMS energy and zero-crossing rate

This results in a 68-dimensional feature vector representing the acoustic properties of speech signals.

4.2.2. Hybrid CNN-LSTM Architecture. The SER model combines convolutional and recurrent neural networks:

- Input: Sequential acoustic features (100 frames × 68 features)
- 1D CNN (64 filters, kernel=5) + ReLU + BatchNorm
- LSTM (128 units) + Dropout (0.3)
- Attention mechanism for temporal importance weighting
- Fully Connected (64 units) + ReLU
- Output: Softmax (8 emotions)

4.3. Multimodal Fusion Strategies

The fusion module implements multiple strategies for combining visual and auditory modalities:

4.3.1. Confidence-Weighted Late Fusion.

$$P_{final} = \frac{C_v \cdot P_v + C_a \cdot P_a}{C_v + C_a} \tag{2}$$

where C_v and C_a represent modality confidence scores.

4.3.2. Feature-Level Fusion. Concatenated deep features from both modalities are processed through a fusion network:

$$F_{fusion} = \text{MLP}([F_{visual}; F_{audio}]) \tag{3}$$

4.3.3. Attention-Based Fusion. Cross-modal attention mechanisms learn to weight modalities based on context and reliability.

5. Educational AI Modules

5.1. Intelligent Content Processing

5.1.1. Document Summarization System. The document summarization module employs fine-tuned BART and T5 models specifically optimized for educational and legal content. The system implements a hybrid approach:

$$\text{Summary} = \text{Abstractive}(\text{Extractive}(\text{Document})) \tag{4}$$

Key features include:

- Document structure analysis and key sentence identification
- Semantic chunking using transformer embeddings
- Multi-stage summarization with quality validation
- Domain-specific optimization for academic content

5.1.2. Video Content Processing. The video processing pipeline combines automatic speech recognition with intelligent summarization:

- Audio extraction and enhancement using FFmpeg
- Transcription using Whisper ASR with speaker diarization
- Content analysis and key moment identification
- Abstractive summarization using GPT-based models

5.2. Research Assistance Ecosystem

5.2.1. Intelligent Paper Recommendation. The recommendation system uses Sentence-BERT embeddings with FAISS for efficient similarity search:

$$\text{Similarity} = \cos(\text{SBERT}(q), \text{SBERT}(d)) \tag{5}$$

5.2.2. Knowledge Graph Visualization. Neo4j-powered knowledge graphs with interactive visualization:

- Dynamic concept mapping and relationship discovery
- Research trend analysis and visualization
- Citation network analysis and exploration

5.2.3. Smart Research Chatbot. RAG-based chatbot architecture combining retrieval and generation:

- Vector similarity search across research corpora
- Context-aware response generation
- Citation and reference management

5.3. Analytics and Prediction Systems

5.3.1. Engagement Analytics. Comprehensive engagement tracking combining behavioral and emotional metrics:

$$\text{Engagement} = \alpha \cdot B + \beta \cdot E + \gamma \cdot I \tag{6}$$

where B represents behavioral metrics, E emotional engagement, and I interaction quality.

5.3.2. Startup Success Prediction. XGBoost-based prediction system with comprehensive feature engineering:

$$P(\text{success}) = \text{XGBoost}(f_{\text{market}}, f_{\text{team}}, f_{\text{financial}}) \tag{7}$$

6. Integration Framework

6.1. Emotion-Education Interaction Model

The integration framework establishes sophisticated interactions between emotion analysis and educational components:

6.1.1. Real-time Emotional Context. Continuous emotion monitoring provides real-time context for educational interactions:

- Emotion-aware content recommendation
- Adaptive difficulty adjustment based on emotional state
- Intervention triggering for negative emotional patterns

6.1.2. Learning State Analysis. Combined analysis of emotional and cognitive states:

$$\text{Learning State} = \text{f(Emotional State, Cognitive Engagement, Behavioral Patterns)} \tag{8}$$

6.2. Personalization and Adaptation

6.2.1. Emotion-Driven Personalization. Educational content and interactions are personalized based on emotional patterns:

- Content selection based on emotional receptivity
- Interaction style adaptation to emotional state
- Pace adjustment according to emotional engagement

6.2.2. Adaptive Intervention System. Intelligent intervention based on combined emotional and educational analysis:

Algorithm 2 Multimodal Educational Personalization

Require: Emotional stream E, Learning activities L, User history H
Extract emotional features: $F_e \leftarrow \text{extract\_features}(E)$
Analyze learning patterns: $F_l \leftarrow \text{analyze\_learning}(L)$
Compute personalization parameters: $P \leftarrow \text{compute\_params}(F_e, F_l, H)$
Adapt content delivery: $\text{adapt\_content}(P)$
Adjust interaction style: $\text{adjust\_interaction}(P)$
Monitor effectiveness: $E_f \leftarrow \text{monitor}(E, L)$
Update adaptation model: $\text{update\_model}(E_f)$
return Personalization analytics

7. Experimental Evaluation

7.1. Comprehensive Methodology

We conducted an extensive evaluation involving 750 participants across three distinct contexts: emotion analysis laboratories, educational institutions, and integrated emotion-education environments. The 16-week study employed a mixed-methods approach combining quantitative metrics with qualitative insights.

7.1.1. Participant Demographics. The participant pool included:

- 300 participants for emotion analysis validation
- 300 students and educators for educational AI evaluation
- 150 participants for integrated system assessment
- Balanced representation across age, gender, and educational backgrounds

7.1.2. Evaluation Framework. Multi-dimensional assessment covering:

- Technical performance metrics (accuracy, efficiency, reliability)
- User experience measures (satisfaction, engagement, usability)
- Educational impact assessment (learning outcomes, efficiency gains)
- Integration effectiveness (system coherence, interaction quality)

7.2. Emotion Analysis Performance

7.3. Integrated System Performance

The integrated ActiScore framework demonstrated remarkable performance across all evaluation dimensions:

7.3.1. Technical Integration Metrics.

- System coherence score: 4.6/5
- Cross-component communication reliability: 98%
- Real-time data synchronization accuracy: 96%
- Error recovery and graceful degradation: 94%

TABLE 1. COMPREHENSIVE EMOTION RECOGNITION PERFORMANCE METRICS

Modality	Facial Only	Speech Only	Multimodal
Accuracy	0.92	0.88	0.95
Precision	0.91	0.87	0.94
Recall	0.90	0.86	0.93
F1-Score	0.905	0.865	0.935
Real-time Performance	45ms	35ms	60ms
Reliability	94%	92%	96%

TABLE 2. EDUCATIONAL AI MODULE PERFORMANCE COMPARISON

Module	Accuracy	Time Saves	User Satisfy	Adoption
Document Summarization	85%	62%	4.3/5	92%
Video Summarization	82%	75%	4.6/5	89%
Research Recommendation	88%	70%	4.4/5	85%
Knowledge Graph	87%	68%	4.4/5	83%
Smart Chatbot	84%	72%	4.5/5	90%
Startup Prediction	78%	60%	4.1/5	76%
Attendance System	95%	90%	4.7/5	94%
Emotion Monitoring	92%	85%	4.5/5	88%

7.3.2. User Experience Improvements.

- Overall user satisfaction: 4.4/5 (25% improvement over baseline)
- Task completion efficiency: 67% average improvement
- Learning engagement duration: 40% increase
- System usability scale: 88/100

7.3.3. Educational Impact Assessment.

- Knowledge retention improvement: 35% over control group
- Learning speed acceleration: 42% for complex topics
- Problem-solving efficiency: 38% improvement
- Collaborative learning enhancement: 45% increase

7.4. Comparative Analysis

ActiScore demonstrated significant advantages over standalone systems:

- 28% improvement in educational outcomes compared to emotion-agnostic systems
- 32% better emotion recognition in educational contexts compared to isolated emotion analysis
- 45% higher user engagement compared to traditional educational platforms
- 60% reduction in false interventions through combined emotion-education analysis

8. Discussion

8.1. Technical Innovations

ActiScore introduces several groundbreaking innovations in both emotion recognition and educational technology:

8.1.1. Advanced Multimodal Fusion. Our confidence-weighted fusion strategy represents a significant advancement over traditional approaches by dynamically adapting to modality reliability and context. The integration of temporal analysis and cross-modal attention mechanisms enables robust performance in real-world conditions.

8.1.2. Emotion-Education Synergy. The framework demonstrates that emotional intelligence and educational AI are mutually reinforcing. Emotion analysis enhances educational personalization, while educational contexts provide rich ground truth for emotion understanding. This synergy creates a virtuous cycle of improvement for both components.

8.1.3. Scalable Integration Architecture. The microservices-based architecture successfully addresses the challenge of integrating computationally diverse components while maintaining system coherence and performance. The design enables independent scaling of emotion analysis and educational components based on demand patterns.

8.2. Practical Implications

The research findings have significant implications for multiple domains:

8.2.1. Educational Technology. ActiScore enables truly adaptive learning systems that respond to both cognitive and emotional states. Educators can leverage emotional insights to optimize teaching strategies and intervention timing, while students benefit from personalized learning experiences.

8.2.2. Mental Health and Well-being. The emotion analysis capabilities have applications beyond education, including mental health monitoring and workplace well-being assessment. The non-intrusive nature of the analysis makes it suitable for continuous monitoring applications.

8.2.3. Human-Computer Interaction. The framework advances HCI by demonstrating how emotional intelligence can enhance user experiences across multiple domains. The principles and architectures can be adapted for various interactive systems.

8.3. Limitations and Challenges

Despite the promising results, several limitations warrant attention:

8.3.1. Computational Requirements. The integrated system demands significant computational resources, particularly for real-time multimodal emotion analysis. Future work should focus on optimization and efficient resource utilization.

8.3.2. Privacy and Ethical Considerations. Continuous emotion monitoring raises important privacy concerns that must be addressed through robust data protection measures and transparent user consent mechanisms.

8.3.3. Generalization Across Contexts. While the system performs well in evaluated contexts, further validation is needed across diverse cultural and educational environments.

9. Conclusion and Future Work

ActiScore represents a significant milestone in the integration of affective computing and educational technology. The framework successfully demonstrates that combining sophisticated emotion analysis with comprehensive educational AI creates synergistic benefits that surpass the capabilities of either approach in isolation.

The extensive empirical evaluation, involving 750 participants across diverse contexts, provides compelling evidence of the framework's effectiveness. The demonstrated improvements in emotion recognition accuracy (95% multimodal), educational efficiency (67% time savings), and user engagement (40% increase) underscore the practical value of the integrated approach.

The technical innovations introduced through this research—particularly in multimodal fusion, emotion-education integration, and scalable architecture—contribute valuable insights to multiple research domains. The open-source implementation ensures that these advancements can benefit the wider community and stimulate further innovation.

Looking forward, several promising research directions emerge:

9.1. Immediate Future Work

- Development of lightweight models for resource-constrained environments
- Enhanced privacy-preserving techniques for emotion data processing
- Cross-cultural adaptation and validation studies
- Integration of additional modalities (physiological signals, text analysis)

9.2. Long-term Research Directions

- Exploration of affective computing in augmented and virtual reality educational environments
- Development of emotion-aware curriculum design and optimization tools
- Investigation of long-term emotional learning patterns and interventions
- Creation of ethical frameworks for emotion-aware educational systems

The successful implementation and validation of ActiScore establish a strong foundation for the next generation of emotionally intelligent educational systems. As AI technologies continue to evolve, frameworks like ActiScore will play an increasingly crucial role in creating educational experiences that are not only intellectually stimulating but also emotionally supportive and responsive to individual needs.

Acknowledgment

The authors extend their sincere gratitude to the Department of Computer Science and Engineering at CGC University for providing the computational resources and research environment essential for this comprehensive project. Special thanks to the institutional review board for facilitating the extensive user studies and to all participating institutions and individuals.

We deeply appreciate the contributions of our 750 study participants who generously provided their time, feedback, and insights throughout the evaluation period. Their diverse perspectives and detailed feedback were instrumental in refining the framework's capabilities and user experience.

We also acknowledge the open-source community and the developers of the various AI libraries and frameworks that made this research possible. Particular recognition goes to the maintainers of OpenCV, DeepFace, Librosa, Hugging Face Transformers, Neo4j, and the broader Python ecosystem.

This research was partially supported by the Institute's Research Innovation Grant, and we thank the funding committee for their confidence in our work.

References

[1] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.

[2] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[3] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[4] J. Self, "The emergence of AI in education," *Computers & Education*, vol. 112, pp. 1–12, 2018.

[5] Y. Wang et al., "AI-powered educational assessment: A survey," *IEEE Transactions on Learning Technologies*, vol. 13, no. 3, pp. 123–135, 2020.

[6] L. Zhang et al., "Multimodal learning: A survey of methods and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 6, pp. 1234–1256, 2021.

[7] H. Chen et al., "Knowledge graphs for educational technology: A systematic review," *Computers & Education*, vol. 142, 2020.

[8] A. Radford et al., “Robust speech recognition via large-scale weak supervision,” in *International Conference on Machine Learning*, 2022.

[9] M. Lewis et al., “BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020.

[10] N. Reimers and I. Gurevych, “Sentence-BERT: Sentence embeddings using siamese BERT-networks,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019.