Affective Computing in Intelligent Tutoring Systems: Exploring Insights and Innovations

Theofanis Tasoulas, Christos Troussas, Phivos Mylonas, Cleo Sgouropoulou
Department of Informatics and Computer Engineering, University of West Attica
Egaleo, Greece
{mscacs22024, ctrouss, mylonasf, csgouro}@uniwa.gr

***Abstract*— Integrating intelligent tutoring systems (ITS) into education has significantly enriched personalized learning experiences for students and educators alike. However, these systems often neglect the critical role of emotions in the learning process. By integrating affective computing, which empowers computers to recognize and respond to emotions, ITS can foster more engaging and impactful learning environments. This paper explores the utilization of affective computing techniques, such as facial expression analysis and voice modulation, to enhance ITS functionality. Case studies and existing systems have been scrutinized to comprehend design decisions, outcomes, and guidelines for effective integration, thereby enhancing learning outcomes and user engagement. Furthermore, this study underscores the necessity of considering emotional aspects in the development and deployment of educational technology to optimize its influence on student learning and well-being. A major conclusion of this research is that integration of affective computing into ITS empowers educators to customize learning experiences to students' emotional states, thereby enhancing educational effectiveness. To achieve this, it is crucial for such systems to make informed design decisions and adhere to established guidelines. These design decisions may involve selecting appropriate emotion recognition algorithms and integrating them seamlessly into the ITS interface, while guidelines can offer practical insights into effectively implementing affective computing techniques in ITS, such as ensuring data privacy and security measures are in place and providing adequate training for educators on interpreting emotional data.**

Keywords*— *Affective Computing; Intelligent Tutoring Systems; Sentiment Analysis; Emotion Detection; Exploratory study

I. Introduction

Affective computing is a branch of computer science and artificial intelligence (AI) that aims to develop software capable of perceiving, interpreting, and responding to human emotions. These technologies incorporate human expressions in data format, such as audio, text and images, to enhance human-computer interaction and improve system intelligence. Affective computing draws from psychology and computer science, involving the collection and processing of data from sources like cameras and microphones. The goal is to build a model of AI that enables personalized computing systems to accurately perceive and manage human emotions. Implementing such systems not only provides insights into human emotion functioning. but also has the potential to enhance emotional intelligence for both software engineers and users. Emotional intelligence, crucial for understanding and managing emotions effectively, shares similarities with the process by which software acquires emotional intelligence—through data collection, processing, and responsive actions [1], [2].

Emotion or sentiment recognition refers to the process of analyzing emotional states expressed through data. Machine learning models, such as neural networks, are trained on datasets to classify data into specific emotions. Models learn to recognize patterns in pre-specified features and associate them with different emotions, such as happiness, sadness, anger, etc. Once the model is trained, it can be used for real-time emotion recognition by continuously analyzing features from data, such as facial expressions from a video, which is imported from a camera. The system can detect and monitor emotions as they change over time, allowing human-computer interaction applications to be upgraded. Emotional recognition utilizes various types of data, including written text, spoken language, facial expressions, and physical signals [1], [2].

ITS use advanced technology, data analysis, and AI in order to provide personalized instruction and manage educational content. These systems can adjust the content based on how students perform, their learning styles, and their pace, improving the learning experience and keeping students engaged. They use multimedia or the internet and give students real-time feedback, while also allowing teachers to assess student performance and adjust their teaching strategies. AI is essential in creating these intelligent training programs by monitoring how students interact and analyzing data like test scores and emotional responses. AI algorithms can identify areas where students need more help and suggest personalized learning approaches, improving learning outcomes and thereby reducing the teacher's direct intervention. Moreover, AI-driven educational tools promote problem-solving skills, accuracy in sentiment assessment, and data-driven decision-making for educators and institutions. In special education, ITS hold potential for supporting students with disabilities and learning difficulties. While this area is relatively unexplored, special education programs typically involve individualized learning plans, assistive technologies, and special adaptive teaching strategies to ensure equal access to quality education and promote academic, social, and emotional development [3], [4]. Especially when it comes to emotion detection, people with special needs are often limited to express their feelings. This is a great challenge, not only for ITS that incorporate emotion detection, like detection of facial expressions, but also for special education teachers. And that is why further investigation should be done, on how to understand human sentiment from individuals' expressions.

In view of the above, this research aims to highlight the affective computing techniques that enhance ITS in order to improve students' learning experience. In addition, the current paper discusses the structure and enhancement of ITS, illustrating how sentiment analysis support students, even when they have disabilities and learning difficulties. Also, by reviewing existing ITS implementations and case studies, this research showcases how affective computing technologies are already being utilized to enhance teaching and learning experiences, while also discussing the challenges and limitations associated with these technologies and suggesting areas for further research and improvement. Overall, this article aims to underscore the importance of considering emotional aspects in the development and deployment of ITS, advocating for the integration of affective computing techniques to optimize experiences and support emotional management in learning.

II. Sentiment detection techniques

A. Face detection

Face detection algorithms aim to detect faces and their features (e.g., nose, mouth, eyes, etc.) in images or videos using computer vision algorithms. It is important that this is done regardless of their position, orientation, age, expression, and lighting conditions, in order to have a great efficiency. Facial recognition technology finds utility in authentication, access control, and emotion detection [5]. This process involves scanning an input image using a sliding window approach, with statistical learning methods used to train classifiers. However, challenges arise from factors such as changes in facial appearance, lighting conditions, expression, and head pose, as mentioned before, requiring nonlinear classifiers to handle complex boundaries. Then image preprocessing techniques should apply, like grayscale conversion, noise reduction, contrast enhancement, normalization, image resizing, etc. After that, for face feature recognition usually geometric and neural methods used. Geometric methods detect facial features based on shapes and compute properties like distances and angles. For instance, detecting two ellipses with two concave-down curves above them in an image, could be translated as two eyes and their eyebrows. Those methods use graphical computation algorithms, while neural methods operate directly on an image-based face representation, analyzing pixel intensity matrices to identify specific facial features [6]. Otherwise, already implemented techniques like Gabor wavelet representation and independent component analysis facilitate effective face recognition [7]. After identifying the facial features, their movements and changes could imply specific emotions. For example, two eyebrows drawing together, could mean that the person is angry. This could be done by distance calculation methods on facial features, including Euclidean Distance, Square Euclidean Distance, and CityBlock Distance. Last but not least, image classification algorithms can detect changes, such as whether a mouth is open or closed or if a person is smiling or not, by comparing an isolated image of a mouth with a pre-trained model [8].

B. Speech analysis

Another way for sentiment recognition lies in speech analysis. However, achieving accurate sentiment recognition remains a significant challenge due to the uncertainty in selecting the right features and the presence of background noise, which can hamper machine learning models' effectiveness. The complexity of recognizing emotions from speech surpasses that of visual signals or written text due to the unique representation of audio signals. Unlike text, audio signals are typically represented as one-dimensional matrices or sequences of binary values encoding acoustic waveform amplitudes at discrete time points. This complex digital representation requires intricate processing, including analog-to-digital conversion [9]. In audio preprocessing, techniques such as sampling, normalization, and noise removal are commonly employed. Subsequently, voice or speech recognition involves identifying spoken content and detecting the user's emotions during speech, based upon audio features, such us the pitch (frequency), the rate of speech (words per time) or the intensity of the voice (loudness). For instance, a fast and high-pitched speech may indicate excitement or happiness, while a slow and monotone speech rate may suggest sadness or boredom. Usually, supervised machine learning algorithms trained on annotated data are used to facilitate those sound features. While manual labeling may lead to variations in accurate emotion detection based on the supervisor's ability to identify the correct sentiment, prior sentiment knowledge can substantially enhance the algorithm's computing performance [10]. Thus, feature extraction captures spectral features and variations to classify emotional states, with interpretations varying based on the researcher or programmer [11]. Combining different types of features enhances robust sentiment recognition, with particular emphasis on pitch and harmony classes. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), either with long-term or short-term memory, or a combination, are commonly employed architectures in speech-based sentiment recognition frameworks [9].

C. Text analysis

Sentiment analysis of text involves extracting opinions and emotions expressed by the author, utilizing data mining and classification techniques. This method could find applications in various fields such as customer satisfaction assessment, e-learning materials enhancement, and mental health disorder prediction. Analyzing attitudes on multiple topics within a single text presents challenges both for algorithms and human comprehension, often requiring individual examination of each sentence or section. Conflicting views within a text or sentence further complicate sentiment analysis, as exemplified by sentences expressing both positive opinions and negative feelings simultaneously [12]. Text can be input via keyboard or converted from handwritten form to digital text through optical character recognition (OCR). OCR algorithms preprocess images by reducing noise, correcting skew, and converting images to binary format. Character recognition, accomplished by using machine learning models like convolutional neural networks, identifies and classifies individual characters based on visual characteristics. The recognized characters are then compiled into computer-readable text, often with the aid of natural language modeling to enhance accuracy and correct errors. Techniques like text normalization, tokenization, and word removal streamline text preprocessing, while word embedding methods like Word2Vec and BERT convert words into numeric vectors, capturing semantic relationships. Moreover, Natural Language Processing (NLP) is a technology that plays a crucial role in understanding, interpreting, and creating human language. It involves syntax and grammar analysis to understand sentence structure. However, sentiment analysis primarily relies on sentiment lexicons and machine learning models to determine text emotion. Those lexicons or dictionaries are pre-defined labeled words or phrases, based on which the emotion classification will be done. This is called sentiment mining, which determines the emotional content of the text. In addition, opinion mining, which focuses on identifying opinions expressed in text, might also help sentiment mining, as well as subjectivity detection, which distinguishes subjective texts expressing personal opinions from objective texts conveying factual information. Generally, sentiment classification methods utilize dictionaries to map terms to emotions, often combining supervised and unsupervised learning techniques for better results. Additionally, punctuation marks and text features can influence sentiment estimation, though syntax and text size typically play minor roles [12]. Sentiment analysis algorithms employ various procedures and features for classification, often relying on supervised learning with labeled data. Descriptive features and dictionary-based approaches are common, where dictionaries provide information about the type and strength of emotion conveyed by words or phrases. Negation handling and punctuation analysis further refine sentiment estimation, while features like syntax and text size have minimal impact on sentiment classification [13].

D. Physical signals analysis

Other physical or physiological signals, such as heart rate, skin conductance and brain activity, may also provide additional information about emotional states. Physical sensors algorithms are used to interpret these signals, usually through the electrical signals passing from the individual to a sensor. Then, simply measuring them can determine an emotional state. For instance, if the heart rate is high, it can imply anxiety. However, there is limited knowledge and research done on those features' analysis. Last but not least, other methods of measuring biological features or substances, such as dopamine, could be useful, but these biomarkers may be difficult to be obtained and identified in a short period of time or even in real time.

E. Models for sentiment classification

Various classifiers, including neural networks, Support Vector Machines (SVMs), and Hidden Markov Models (HMMs), are employed in order to classify the emotions from data, based on extracted features. The emotion recognition methods are broadly categorized into context-based and sequence-based methods. Context-based recognition relies on individual data, often without temporal context, such as a specific image. Techniques such as neural networks and SVMs are commonly used in this approach. In contrast, sequence-based methods leverage temporal information from a sequence of data to recognize expressions, like a sequence of images on a video, utilizing models like HMMs and Recurrent Neural Networks (RNNs) to capture the dynamics of data's features over time [7]. HMMs are powerful statistical tools for sequential data modeling, extensively used in facial expression recognition. HMM-based facial expression recognition involves several steps, including data preprocessing, feature extraction, model definition, and classification. The Baum-Welch algorithm is often used for parameter estimation in HMMs, enabling the model to capture time dependencies and dynamics in emotional changes over time. Despite the rise of deep learning methods, HMMs remain valuable, when dealing with limited data or when the ability to interpret is a priority [14]. RNNs are well-suited for facial expression recognition tasks too, because of their ability to capture temporal dependencies. Backpropagation algorithm through time enables RNNs to learn from temporal sequences, gradually improving performance. Convolutional neural networks (CNNs), on the other hand, are more complex neural networks, automatically learning and representing complex patterns in features of images, video frames or even of audio. By analyzing spectrogram representations of audio data and by capturing localized features, CNNs learn to classify sound-based emotions effectively. Training involves backpropagation and gradient descent to optimize parameters, minimizing the difference between predicted and actual emotions [15], [16]. Moreover, classification and clustering algorithms play crucial roles in sentiment analysis. Naive Bayes, despite its "naive" assumption of feature independence, performs well in sentiment recognition tasks, especially of text analysis, even with large datasets. K-means clustering, an unsupervised method, groups similar data points based on distance measurements. However, the effectiveness of those algorithms may vary depending on the nature of the data and their parameters, e.g., the number of clusters selected in K-means [17]. In conclusion, sentiment classification methods encompass a range of techniques and algorithms tailored to the specific requirements of emotion analysis systems. Context-based and sequence-based methods offer distinct advantages, with researchers exploring various classifiers and evaluation methodologies to enhance accuracy and robustness. Nevertheless, the exact measurements of accuracy and effectiveness often are missing, which is crucial for selecting the right model or algorithm.

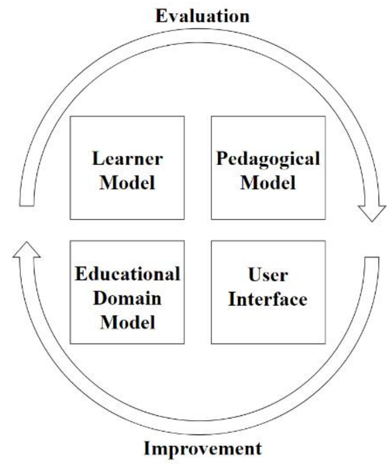

Fig. 1. Architecture of Intelligent Tutoring System

III. INTEGRATING AFFECTIVE COMPUTING INTO INTELLIGENT TEACHING SYSTEMS

A. Structure of intelligent teaching systems

ITS have gained attention especially in academic and school environments, while encompassing various approaches. Web-based systems facilitate online learning through multimedia, like video lectures, interactive exercises, and personalized learning paths, offering flexibility in learning. Interactive systems leverage technologies such as web cameras and speech recognition though microphones, in order to create exciting learning experiences. In addition, they utilize AI and big data technologies to evaluate teaching quality by analyzing various data, such as images and text related to students' performance [18]. The fundamental architecture of an ITS comprises four basic elements: the learner model, the pedagogical model, the educational domain model, and the user interface [19] (Fig. 1). The learner model identifies learner characteristics and preferences, while the pedagogical model applies strategies to deliver content tailored to learner's preferences. The educational domain model encompasses the knowledge and content to be learned, including multimedia, rules, and learning methods. The user interface facilitates smooth interaction between the learner and the system or between individuals within the learning environment [3]. Furthermore, quantitative evaluation methods are used for teaching quality measurement, including sophisticated techniques such as one-way ANOVA, Markov chain analysis, artificial neural networks, and data mining. These methods can be applied in evaluating teaching quality in higher education, even though most researches do not include quantitative evaluation, rather some general qualitative evaluation [18]. Based on the evaluation, changes can be applied for further improvement of the ITS.

B. Enhancement intelligent teaching systems

ITS can leverage technologies such as speech recognition, computer vision, and NLP to enhance human-computer interaction and educational effectiveness. Speech recognition technology enables analysis of teachers' speech during lectures, assessment of pronunciation and fluency in language learning modules. Computer vision technology analyzes student behavior and emotional states in classrooms, providing feedback to enhance teaching quality. NLP technology facilitates effective human-computer interaction, enabling analysis of spoken or written responses and automatic correction of assessments [5]. Combining these technologies can further strengthen the intelligence of educational systems. For example, speech recognition combined with NLP could enable automatic assessment of spoken or written responses to open-ended questions. These integrated approaches offer valuable insights into student learning and engagement, enabling personalized and effective teaching strategies [18]. Obviously affective computing can be integrated to ITS to create a more attractive education environment, as well as an easier method of learning. And that's because when emotions play significant role in teaching environments and managing the student's emotions can further strengthen their learning experience. Education becomes more exciting when individuals are in better emotional states, as they tend to prefer learning under such conditions.

C. Intelligent teaching systems in special education

Intelligent teaching systems designed for special categories of students, such as those with dyslexia or autism spectrum disorder (ASD), require careful consideration of psychological aspects and tailored approaches to address their unique needs. For individuals with dyslexia, an ITS can help in identifying and addressing reading and writing difficulties. NLP models analyze written or spoken language to detect syntactic or semantic errors characteristic from the dyslexic student. Thereby, the system can adapt its approach to accommodate the learner's needs. For instance, it can provide alternative word suggestions for misspelled and track the improvement of the student's dyslexia. Similarly, individuals with low vision benefit from technologies like Microsoft's Visual Assistive Seeing VR Toolkit, which leverages virtual and augmented reality to create immersive learning environments. These technologies enhance accessibility by providing automatic computer reading and virtual mapping for navigation, reducing reliance on traditional visual aids like glasses or contact lenses. Additionally, individuals with mobility difficulties can utilize augmented or virtual reality to engage in more interactive and dynamic learning experiences. In the case of autism spectrum disorder (ASD), ITS can utilize machine learning and NLP algorithms to detect and support learners with ASD. By analyzing language patterns, including repetitive phrases, difficulty with metaphors, and emotional expressiveness, the system can identify potential challenges and tailor interventions accordingly. Moreover, tracking eye gaze patterns and facial expressions during training sessions can provide insights into social interaction difficulties and emotional regulation issues characteristic of ASD. These systems can offer real-time feedback, suggest corrections, and simulate social interactions to help individuals with ASD develop essential social skills in a safe and controlled environment. However, it's essential to acknowledge that these systems rely on predefined algorithms and require ongoing refinement to effectively interpret and respond to individual needs. While technology can enhance accessibility and support for individuals with special needs, human oversight and intervention remain critical to ensure the ethical and effective implementation of intelligent teaching systems in special education contexts [3].

D. Existing intelligent tutoring systems

Existing applications of intelligent teaching systems encompass a diverse range of approaches and technologies aimed at enhancing the educational experience and addressing specific challenges in teaching and learning (Fig. 2). Traditional ITS, such as Andes Physics Software [20], offer an open-ended approach to teaching, allowing students to solve problems in physics using various methods. It employs a solution graph structure to provide immediate feedback and assistance to students. Despite some bugs, Andes Physics Software has been found to improve students' physics grades over years, although it focuses heavily on physics exercises, potentially limiting broader knowledge acquisition. Another system, called ITSPOKE [21], facilitates oral dialogue between students and a computer to solve qualitative physics problems. It integrates speech recognition and NLP to provide feedback and explanations to students. While effective in promoting spontaneous explanations, ITSPOKE faces challenges with speech recognition accuracy. Moving on, Gao B. [18] has created a behavior analysis system, which utilizes cameras and big data training to analyze teacher and student behavior real time in the classroom. It employs computer vision through web cameras and AI software to automate behavior recognition (e.g., students playing on phone, raising hands or sleeping) and assess the quality of education. By detecting behaviors and emotions, the system provides valuable insights for improving teaching strategies and student engagement. Another ITS, implemented by [22]., integrates real-time facial recognition and real time sentiment analysis during lectures. It combines convolutional neural networks with facial recognition technology to provide teachers with insights into students' emotional experiences and responses, enhancing the educational environment. An alternative ITS called AICARP.V2 [23] focuses on detecting changes in users' emotional states through multi-sensory support. This ITS includes physiological sensors and actuators on the student's hand, in order to monitor and mitigate stress levels during learning activities. By integrating real-time data analysis and sensory feedback, AICARP.V2 offers personalized support for students, while performing the educational activities (e.g., English speaking exams). Also, the same issue of personalized support for students is also addressed in [24], where user knowledge has been modeled to personalize the training and learning path, particularly in training engineering students' spatial skills. Similarly, adaptive systems using emerging technologies such as augmented reality have shown great promise in personalizing educational experiences. Last but not least, innovative and automated training systems with robots, such as Pepper [25], are developed to identify students' behaviors, enhancing engagement and interaction for autistic children. The system, equipped with a camera and microphone, assesses children's attention, behavior and emotional expressions in real-time, tailoring interactions based on each child's needs and progress. Thus, autistic children tended to attend more the robot in classroom. The capabilities of the robot Pepper extend beyond traditional teaching methods, by adopting specific postures, adjusting speech volume, and incorporating visual cues like flashing LEDs, in order to adapt to negative emotional states such as "distracted" or "confused". Additionally, the Pepper robot recognizes emotions from facial expressions, engages students by administering multiple-choice quizzes and recites texts from its knowledge base, fostering critical thinking and problem-solving skills. While it typically supports teachers in the classroom under supervision, its ability to respond dynamically to students' emotional cues enhances engagement and facilitates a great improvement of ITS [26].

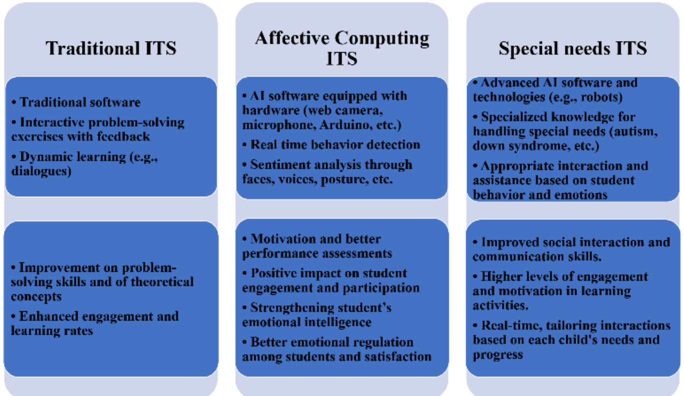

Fig. 2. Types of Iintelligent Tutoring Systems, including their Features and their outcomes.

IV. Discussion

A. Challenges

A major challenge in emotional recognition is the subjectivity and variability of emotions, which may differ across individuals and contexts. Consistently defining emotions is challenging, because of the diverse biases among interpreters and variations in their emotional intelligence. Another challenge lies in extracting relevant features from various data types like text, audio, and facial images. While combining these datasets enhances model efficiency and understanding, it increases complexity as inferences from one dataset influence others. For the moment, the current ITS do not combine lots of sentiment analysis techniques. Emotions can be multifaceted, making categorization difficult, especially when conflicting emotions coexist. Sophisticated models capable of handling overlapping emotional states are needed, assigning probabilities to multiple emotional labels. Recognizing the context of emotional states is crucial for deeper understanding, as emotions are influenced by environmental and personal factors. However, privacy concerns may arise, for sentiment data may be considered sensitive. Implementing privacy policies and robust data security measures is essential to ensure proper management of students' data. Furthermore, learning methods should be implemented, accommodating various learning preferences and fostering a supportive learning environment. Despite challenges, advances in AI and machine learning are driving progress in affective computing, as it remains a dynamic and evolving field, continuously adapting to new developments and applications.

B. Limitations of our research

The current research offers valuable insights, but also faces several limitations. While different methods and instances are presented, they frequently concentrate on specific case studies, which restricts the applicability of results to broader educational settings. Moreover, many studies are recent, leaving uncertainty about the long-term effectiveness and impact of these systems on education. Especially when it comes on measuring the effectiveness, most researchers do not include quantitate measurements of the sentiment classification and the ITS itself (e.g., precision, accuracy, etc.). Also, ethical issues are not discussed, including the privacy policies, the change of the role of teachers in classrooms, the assurance of equal access to quality education, the biases on the pre-trained models. Finally, the present work focuses on basic emotions and expressions, but recognizing a wider range of emotions and their context can be a challenge for these technologies. Addressing these limitations will be critical to unlock the potential of ITS, for they require substantial improvements to ensure effectiveness and emotional engagement.

V. Conclusions

In summary, this paper provides an overview of integrating affective computing methods and examples in ITS. Using sentiment analysis students can further develop their knowledge into a more attractive learning environment. Moreover, students' personal difficulties or special needs will be assisted by those systems through personalized educational experiences. This requires sentiment extraction from features, such as faces or audio. Those techniques were deeply analyzed for further improvement of the current algorithms. Each algorithm offers unique advantages and hopefully in the future, we will gain greater clarity on which algorithms and techniques are most effective for facial expression, audio, and text analysis. While ITS are equipped with varied capabilities, including interactive activities and sentiment classification, they show potential for even better learning experiences and supporting emotional management. However, development is still in early stages and more research is needed in machine learning, AI and affective computing methods, so that their full potential will be unlocked.

The most significant conclusion of this research is that the ITS have great potentials for strengthening the educational experiences, even though there remains a substantial amount of work yet to be accomplished. More specifically, this paper highlights the limitations of the evaluation system and identifies a scarcity of ITS incorporating combined sentiment analysis techniques. Furthermore, as one delves into the techniques of affective computing, it becomes evident that there is a growing demand for additional benefits derived from the personalized educational experiences offered by these systems. Therefore, the ability to understand and respond to human emotions could be pivotal in the realm of education.

References

[1] R. W. Picard, "Affective computing: challenges," *International journal of human-computer studies*, Jul. 01, 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581903000521

[2] J. Tao and T. Tan, "Affective Computing: A Review," *Lecture notes in computer science*, Jan. 01, 2005. [Online]. Available: https://link.springer.com/chapter/10.1007/11573548_125

[3] S. Dutt, N. J. Ahuja, and M. Kumar, "An intelligent tutoring system architecture based on fuzzy neural network (FNN) for special education of learning disabled learners," *Education and information technologies*, Aug. 27, 2021. [Online]. Available: https://link.springer.com/article/10.1007/s10639-021-10713-x

[4] Z. Sun, M. Anbarasan, and D. Praveen Kumar, "Design of online intelligent English teaching platform based on artificial intelligence techniques," *Comput. Intell.*, vol. 37, no. 3, pp. 1166–1180, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/coin.12351

[5] M. Imani and G. A. Montazer, "A survey of emotion recognition methods with emphasis on E-Learning environments," *Journal of network and computer applications*, Dec. 01, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804519302759

[6] M. O. Faruqe and M. A. M. Hasan, "Face recognition using PCA and SVM," in 2009 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication, 2009, pp. 97–101. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/5276938/

[7] Y.-L. Tian, T. Kanade, and J. F. Cohn, "Facial Expression Analysis," *Springer eBooks*, Nov. 23, 2005. [Online]. Available: https://link.springer.com/chapter/10.1007/0-387-27257-7_12

[8] E. I. Abbas, M. E. Safi, and K. S. Rijab, "Face recognition rate using different classifier methods based on PCA," in *2017 International Conference on Current Research in Computer Science and Information Technology (ICCIT)*, 2017, pp. 37–40. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7965559/

[9] L. Cai, Y. Hu, J. Dong, and S. Zhou, "Audio-Textual Emotion Recognition Based on Improved Neural Networks," Mathematical problems in engineering, Dec. 31, 2019. [Online]. Available: https://www.hindawi.com/journals/mpe/2019/2593036/abs/

[10] M. Papakostas et al., "Deep Visual Attributes vs. Hand-Crafted Audio Features on Multidomain Speech Emotion Recognition," Computation, Jun. 01, 2017. [Online]. Available: https://www.mdpi.com/2079-3197/5/2/26

[11] T. Athanasiadis, P. Mylonas, Y. Avrithis, and S. Kollias, "Semantic image segmentation and object labeling," IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 3, pp. 298–312, 2007. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4118230/

[12] B. Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge, England: Cambridge University Press, 2020. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=v0UBEAAAQBAJ&oi=fnd&pg=PR11&dq=%5B12%5D%09Liu,+Bing.+Sentiment+analysis:+Mining+opinions,+sentiments,+and+emotions.+Cambridge+university+press,+2020.&ots=UgE1z39HuP&sig=KidCKIgH4MsQ08k8xNR8slLbuQE

[13] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, "Current State of Text Sentiment Analysis from Opinion to Emotion Mining," *ACM computing surveys*, May 25, 2017. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3057270

[14] A. Konar and A. Chakraborty, *Emotion recognition: A Pattern Analysis Approach*. John Wiley & Sons, 2015. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=WeOfBQAAQBAJ&oi=fnd&pg=PA147&dq=%5B14%5D%09Konar,+Amit,+and+Aruna+Chakraborty.+Emotion+recognition:+A+pattern+analysis+approach.+John+Wiley+%26+Sons,+2015.&ots=B77ywrGtYe&sig=Xw7toPRlDi8BT8J5dXNguQsP2us

[15] P. Dhruv and S. Naskar, "Image Classification Using Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN): A Review," *Advances in intelligent systems and computing*, Jan. 01, 2020. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-15-1884-3_34

[16] W. Lim, D. Jang, and T. Lee, "Speech emotion recognition using convolutional and Recurrent Neural Networks," in *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2016, pp. 1–4. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7820699/

[17] S. Thavareesan and S. Mahesan, "Sentiment Analysis in Tamil Texts using k-means and k-nearest neighbour," in 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS), 2021, pp. 48–53. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9605839/

[18] B. Gao, "Research and Implementation of Intelligent Evaluation System of Teaching Quality in Universities Based on Artificial Intelligence Neural Network Model," *Mathematical problems in engineering*, Mar. 16, 2022. [Online]. Available: https://www.hindawi.com/journals/mpe/2022/8224184/

[19] C. Troussas, A. Krouska, P. Mylonas, and C. Sgouropoulou, "Personalized learner assistance through dynamic adaptation of chatbot using fuzzy logic knowledge modeling," in *2023 18th International Workshop on Semantic and Social Media Adaptation & Personalization (SMAP)18th International Workshop on Semantic and Social Media Adaptation & Personalization (SMAP 2023)*, 2023, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10255169/

[20] M. Kudělka, J. Pokorný, V. Snášel, and A. Abraham, Eds., *Proceedings of the third international conference on intelligent human computer interaction (IHCI 2011), Prague, Czech Republic, august, 2011*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. [Online]. Available: https://link.springer.com/content/pdf/10.1007/978-3-642-31603-6.pdf

[21] D. J. Litman and S. Silliman, "ITSPOKE: An intelligent tutoring spoken dialogue system," Aclanthology.org. [Online]. Available: https://aclanthology.org/N04-3002.pdf. [Accessed: 20-May-2024].

[22] G. Gavriilidis, "Sentiment analysis during distance learning in an android application and sending it to a RESTful web service". [Online]. Available:

[1] https://search.proquest.com/openview/965018ddbae14efb36ba725e5f8e5fbc/1?pq-origsite=gscholar&cbl=2026366&diss=y (accessed Apr. 19, 2024).

[23] O. C. Santos, R. Uria-Rivas, M. C. Rodriguez-Sanchez, and J. G. Boticario, "An open sensing and acting platform for context-aware affective support in ambient intelligent educational settings," *IEEE Sens. J.*, vol. 16, no. 10, pp. 3865–3874, 2016. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7425146/

[24] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, "Personalization of the Learning Path within an Augmented Reality Spatial Ability Training Application Based on Fuzzy Weights", Sensors 22, 7059, 2022. https://doi.org/10.3390/s22187059

[25] D. Cohen, "A humanoid robot as a therapeutic mediator for young people with autism," *Aldebaran.com*, 2021 [Online]. Available: https://www.aldebaran.com/sites/default/files/blog/2021_NAO_Autism_EN.pdf. [Accessed: 20-May-2024]

[26] Y. Shi, Y. Chen, L. R. Ardila, G. Venture, and M.-L. Bourguet, "A Visual Sensing Platform for Robot Teachers," Sep. 25, 2019. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3349537.3352764