

EMOTION-DRIVEN ADAPTIVE LEARNING SYSTEM FOR REAL-TIME VOICE AND FACIAL EXPRESSION-BASED USER MOTIVATION AND SUPPORT

PROJECT PHASE-1 REPORT

Submitted by

LAVANYA A

210701132

MANISHA SHARMI M

210701146

in partial fulfillment of the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING



**RAJALAKSHMI
ENGINEERING COLLEGE**

An AUTONOMOUS Institution
Affiliated to ANNA UNIVERSITY, Chennai



RAJALAKSHMI ENGINEERING COLLEGE

ANNA UNIVERSITY, CHENNAI

NOVEMBER 2024

ANNA UNIVERSITY CHENNAI - 600 025

BONAFIDE CERTIFICATE

Certified that this project report "**EMOTION-DRIVEN ADAPTIVE LEARNING SYSTEM FOR REAL-TIME VOICE AND FACIAL EXPRESSION-BASED USER MOTIVATION AND SUPPORT**" is the Bonafide work of "**LAVANYA A (210701132), MANISHA SHARMI M (210701146)**" who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Dr. P. Kumar, M.E, Ph.D.,

HEAD OF THE DEPARTMENT

Professor,

Department of Computer Science and

Engineering,

Rajalakshmi Engineering College,

Thandalam,

Chennai-602 105.

SIGNATURE

Mrs. Mahesmeena.K, M.E.,

SUPERVISOR

Assistant Professor,

Department of Computer Science and

Engineering,

Rajalakshmi Engineering College,

Thandalam,

Chennai-602 105.

Submitted to Project Viva-Voce Examination held on_____.

Internal Examiner

External Examiner

ABSTRACT

The proposed e-learning platform introduces AI that can be combined with voice and facial analysis to create a personalized learning experience. By analyzing subtle differences in the user's voice, such as pitch and tone, the system captures emotions ranging from happiness, fear, disgust, happy, sad and stress, making them instantly visible to the learner's heart. In addition, facial recognition technology analyzes subtle changes in the user's face, such as frowning or smiling, to improve the recognition of thoughts. These two types of emotional intelligence allow the system to not rely on any data, but to better understand how the user is feeling at that moment. The platform can adjust its approach when it sees signs of stress, confusion, or conflict, providing support, correcting difficult content, or suggesting a moment to help people learn again. Similarly, if the system detects positive emotions such as satisfaction or confidence, it will push the user to perform higher tasks or provide motivational instructions to stay active and if it detects the user to be sad, it provides a tailor-made motivational or strategic response. Seamlessly integrating sentiment analysis into the learning process creates a more dynamic and supportive environment that meets the needs of learners in ways that traditional platforms cannot. The system not only improves learning outcomes through the development of emotional relationships, but also enhances the user's well-being, making the entire education more holistic and human-centered.

ACKNOWLEDGMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavor to put forth this report. Our sincere thanks to our Chairman **Mr. S. MEGANATHAN, B.E, F.I.E.,** our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN, Ph.D.,** our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** for providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution. Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. P. KUMAR, M.E., Ph.D.,** Professor and Head of the Department of Computer Science and Engineering and our internal guide **Mrs. K. MAHESMEENA, M.E.,** for his guidance and encouragement throughout the project work. We are very glad to thank our Project Coordinator, **Dr.T. KUMARAGURUBARAN, B.E., MTech., Ph.D.,** Department of Computer Science and Engineering for his useful tips during our review to build our project.

LAVANYA A 210701132

MANISHA SHARMI M 210701146

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE
	ABSTRACT	3
	ACKNOWLEDGEMENT	4
	LIST OF FIGS	7
	LIST OF ABBREVIATION	8
1.	INTRODUCTION	9
	1.1 GENERAL	9
	1.2 OBJECTIVE	11
	1.3 EXISTING SYSTEM	12
	1.4 PROPOSED SYSTEM	14
2.	LITERATURE SURVEY	17
3.	SYSTEM DESIGN	21
	3.1 SYSTEM FLOW DIAGRAM	21
	3.2 SEQUENCE DIAGRAM	22
	3.3 CLASS DIAGRAM	23
	3.4 USE CASE DIAGRAM	24
	3.5 ARCHITECHTURE DIAGRAM	25
	3.6 ACTIVITY DIAGRAM	26
	3.7 DATA FLOW DIAGRAM	27

4.	PROJECT DESCRIPTION	28
	4.1 METHODOLOGIES	28
	4.1.1 DATA PREPOCESSING	28
	4.1.2 FEATURE EXTRACTION	29
	4.1.3 CNN AND LSTM ALGORITHM	29
	4.1.4 INCEPTION V3 ALGORITHM	30
	4.1.5 DATA ANALYSIS WITH VOICE	31
	4.1.6 DATA ANALYSIS WITH FACE	31
	4.2 RESULT DISCUSSION	33
5.	CONCLUSION	38
6.	FOR PHASE II	44
7.	REFERENCE	45

LIST OF FIGURES

FIG NO	TITLE	PAGE NO
1	PROPOSED SYSTEM	16
2	PROPOSED SYSTEM	16
3	SYSTEM FLOW	21
4	SEQUENCE DIAGRAM	22
5	USECASE DIAGRAM	23
6	CLASS DIAGRAM	24
7	ARCHETECTURE DIAGRAM	25
8	ACTIVITY DIAGRAM	26
9	DATAFLOW DIAGRAM	27
10	CONFUSION MATRIX	37

LIST OF ABBREVIATIONS

SNO	ABBREVIATION	EXPANSION
1	AI	Artificial Intelligence
2	API	Application Programming Interface
3	AR6	Sixth Assessment Report (IPCC)
4	CNN	Convolutional Neural Network
5	GPT	Generative Pre-trained Transformer
6	IMCC	International Marine Conservation Congress
7	IPCC	Intergovernmental Panel on Climate Change
8	LLM	Large Language Model
9	NLP	Natural Language Processing
10	NASA	National Aeronautics and Space Administration
11	NER	Named Entity Recognition
12	OPT	Open Pre-trained Transformer
13	RAG	Retrieval-Augmented Generation
14	TTS	Text-to-Speech
15	STT	Speech-to-Text
16	DERA	Dialog-Enabled Resolving Agents

CHAPTER 1

1. INTRODUCTION

1.1 GENERAL

The innovative e-learning platform integrates new features, combining exciting analytics and facial recognition to provide on-demand learning, making it a great tool for educational and interactive. These features go beyond the e-learning environment, allowing the system to respond similarly to the user's request by using voice feedback and facial expressions to gauge the user's emotional state. The platform uses advanced techniques to analyze the user's tone, voice, and pitch to identify various emotions, including happiness, sadness, anger, fear, and disgust. It also uses facial recognition technology to detect and interpret changes in facial expressions, further enhancing the accuracy of the process in character analysis.

As users speak or interact on the platform, the system processes both auditory and visual information by combining visual cues from the user's voice with facial cues to understand more of their hearts. For example, if the system detects signs of sadness or stress in the user's voice, or signs of stress or fatigue on their face, the platform will adjust its message response accordingly. It can suggest ways to change negative thoughts, provide words of encouragement, or provide specific resources designed to boost the client's mood and confidence. These suggestions are designed to help users manage their emotions, stay focused, and ultimately improve their learning.

The platform is also designed to detect subtle changes, such as stress or happiness, and provide guidance based on these emotions. If the system detects signs of stress, it will provide warm-up suggestions or remind users to rest. Conversely, if users show excitement or interest, the platforms can provide challenging or motivating content to keep users engaged and maintain a positive learning experience. By integrating emotional intelligence into the learning process, the platform can create a more flexible and effective learning experience.

The combination of audio and visual perception allows the system to provide instant feedback, emotional awareness, and a dynamic that is internalized by users. This emotional intelligence improves the overall user experience as the platform becomes more flexible to students' needs and provides support for people and emotions. This level of emotional awareness creates a sense of unity and support, making users happy and motivated.

Ultimately, this emotional intelligence has the potential to revolutionize online learning by making it more personal and emotional. It allows the system to recognize when users are feeling confused, overwhelmed, or simply need support with the content, and provide instant feedback to help them move on from these thoughts. By creating an environment of knowledge and reflection, the platform not only promotes academic success, but also encourages positive thinking, allowing students to be more motivated, confident, and empowered throughout their education. This partnership helps increase retention, improve overall performance, and create better learning outcomes for students at all levels.

1.2 OBJECTIVE

The aim of the project is to create a smart e-learning platform that aims to increase user engagement and optimize the entire learning process by analyzing and responding to users' emotions. The main goal is to integrate voice and face analysis with cutting-edge technology, allowing the system to instantly detect important emotions such as happiness, sadness, anger, fear, and disgust. By combining learning models for voice recognition and face recognition algorithms, the platform will be able to interpret voice and face cues to evaluate each person's perception of the interaction with the user.

This two-way approach (using both voice and facial analysis) provides a better understanding of the user's emotional state, allowing the platform to respond in a way that supports the learner. For example, if the system detects anxiety or confusion in the user's tone of voice or face, it can provide guidance to calm down, provide additional explanations, or ask them to relax to reduce stress. Similarly, if the system recognizes a positive emotion, such as happiness or joy, it can provide challenging content or additional support to maintain motivation. The thinking skills built into the body allow it to not only passively feel pain, but also proactively demystify the mind without getting in the way of learning. Using this information to personalize the experience, the platform will provide personalized advice, support, and recommendations directly related to the user's emotional state. Whether the user is feeling anxious, confident, worried or passionate, the system can adapt its response to create a nurturing and supportive environment that encourages emotional and good learning. This change will allow students to feel understood and supported throughout their learning, helping to build confidence, reduce stress and ultimately improve learning.

The ultimate goal is to create interactive and responsive e-learning programs that adapt to learners' needs. By integrating knowledge into the platform, the goal is to create a supportive and supportive learning environment, ultimately increasing engagement and learning outcomes.

1.3 EXISTING SYSTEM

The concept of emotional recognition through voice recognition and facial recognition is being explored in the evolving field of educational technology. Some e-learning programs incorporate cognitive-behavioral techniques to increase user engagement and provide more personalized lessons. These platforms typically use a combination of machine learning, such as convolutional neural networks (CNN) to extract features from speech and faces, and short-term temporal networks (LSTM) to identify body patterns in speech. This method helps identify many emotions, such as happiness, sadness, anger, fear, and surprise. Algorithms such as Inception V3 for facial analysis are often used to analyze emotions based on visual cues to gain a deeper understanding of the user's emotional state.

In addition to visualization, language processing (NLP) and speech recognition are also used to better understand the emotional content of a user's communication. This technology allows the system to detect nuances in speech and tone of voice, improving its ability to identify emotions and adjust the response accordingly. By combining these methods, existing systems can adjust their behavior to better match the learner's emotional state, providing personalized feedback such as encouraging more conversation or encouraging when sadness or stress is detected. This makes learning more interactive, responsive, and exciting.

DISADVANTAGES:

Despite these advances, current research methods in e-learning still face many challenges that limit their effectiveness and individual actions:

1. Responsiveness: Many platforms cannot detect or respond to users' emotions in real time. Therefore, learning is less personal, there is no opportunity to interact with students based on their thoughts.
2. General counseling: Although counseling is available, it is often not enough to meet the needs of the individual. General advice or support will not solve the problem of thinking, but will reduce user motivation and engagement. For example, anxious students may not benefit from praise; they may need special guidance or motivation.
3. Limited input: Current systems will rely heavily on visual or textual input, limiting their ability to use all the emotional information available through spoken words and faces. By missing out on instant voice and facial analysis, these platforms fail to capture emotional content, leading to negative emotional evaluations and negative reactions.

Overall, while current platforms have made progress in integrating emotional intelligence, they often fall short in providing personalized learning and good thinking. Not being able to respond to users' emotions immediately reduces their ability to improve coordination, retention, and overall learning.

1.4 PROPOSED SYSTEM

The proposed system further enhances the e-learning process by integrating voice analysis and facial recognition, creating a more interactive and emotional response. Using a combination of neural network (CNN) and short-term (LSTM) networks, the system is able to capture and interpret the emotions in the user's voice in time. While CNNs are used to extract relevant features such as pitch, pitch, and rhythm from the audio input, LSTMs process speech patterns, allowing the system to see changes over time. This enables the platform to analyze the emotions in the mind by analyzing the nuances of the moment of speech and the words spoken.

In addition to speech analysis, the system also integrates facial recognition using the Inception V3 algorithm. This algorithm adds a visual dimension to understanding thoughts by allowing facial expressions to indicate important emotions such as happiness, sadness, anger, fear, and surprise. By combining these two modes (voice and facial expressions), the system improves its ability to interpret the emotional content of user interactions. Binary analysis provides a better understanding of the user's emotional state, which is important for providing timely and relevant responses.

The system also combines speech recognition and natural language learning (NLP) technology to deeply understand the user's emotions and connections to words. NLP helps the system understand the meaning of the user's words and sentences, while speech recognition provides insight by measuring speech patterns and intonation. Together, these technologies allow the platform to provide not only relevant but also emotional feedback and recommendations,

ensuring users feel supported and empowered throughout their learning. This combination leads to a more engaging, personal, and emotional learning experience, ultimately leading to better outcomes.

ADVANTAGES:

1. Self-awareness: The system instantly detects and responds to the user's emotional state, providing self-awareness and encouraging learning. By knowing the emotional content behind the user's words and facial expressions, the platform can adjust its interactions to meet their current needs.
2. Feedback and guidance: The system provides advice, encouragement and guidance based on voice and facial analysis. For example, if it detects that the user is depressed or anxious, the platform will provide positive advice, motivational tips or learning resources to help manage emotions and engage the learner.
3. Instant feedback: Using speech analysis, the system can provide instant feedback on the user's mood and speaking style. This can help users instantly adjust their tone of voice, speaking style, or confidence, which is especially useful for students looking to improve their communication skills (e.g. language learning or public speaking).
4. Increase engagement and motivation: The system increases user engagement and motivation by understanding and acting on users' emotions.

This emotional intelligence can lead to a more productive, responsive, and supportive e-learning environment. Relevant lessons. Serving the emotional and intellectual needs of students, it makes e-learning more interactive, stimulating, and ultimately more effective.



Fig 1 Web interface for voice analysis

The web interface of the Emotion-Driven Adaptive Learning System for voice detection allows users to interact via spoken input. It captures voice data, sends it to the emotion recognition module, and displays the detected emotional state in real time. The interface provides feedback, adjusting content based on emotional cues.

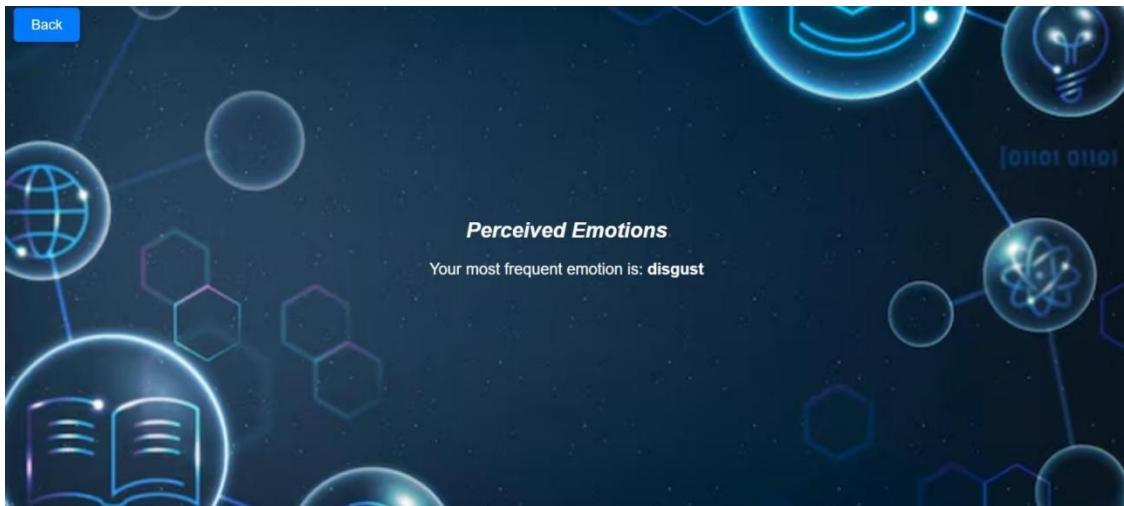


Fig 2 Emotion detection output

The Emotion Detected Page in the Emotion-Driven Adaptive Learning System displays the real-time emotional state of the user based on voice and facial expression analysis. It shows visual indicators of emotions (e.g., happy, stressed, neutral) along with feedback suggestions, such as motivational messages or content adjustments, enhancing user engagement.

CHAPTER 2

LITERATURE SURVEY

1. The research paper titled Audio and Text Sentiment Analysis of Radio Broadcasts (2023) by Naman Dhariwal, Sri Chander Akunuri, Shivama K. and Sharmila Banu explores the potential of sentiment applied to audio data with a focus on radio broadcasting. The research proposes a novel algorithm called forking and blending that combines the best language processing techniques with audio analysis tools like Vokaturi, authoring services like AssemblyAI, and translation services like VADER. This method can extract and classify sentiments from broadcast audio and text, thereby building more effective and real-time opinion models. The study demonstrates the use of this model to analyze sentiments across All India News Press (Akashwani) broadcasts, revealing patterns that indicate changes in public opinion and sentiment. These findings contribute to the advancement of sentiment analysis in the media industry, provide new insights into audience sentiment and public opinion, and lay the foundation for designing more practical applications of opinion analysis in media and communication.
2. Cao Yujing and Pu Jinwan's paper "Music Emotion and Visual Analysis, Integrating Long-Short-Term Memory Network under the Internet of Things" (2023) focuses on the development of automatic analysis of audio emotion music, helping users better understand and understand music. Emotions in music. This paper examines the limitations of traditional music analysis methods and proposes the development of Internet of Things (IoT) technology to overcome these limitations. It uses short-term memory (LSTM) networks to model time series data in music, and combines it with sequence-to-sequence (STS) techniques to build advanced models for music analysis. In the research, various

machine learning algorithms were used to train and evaluate the model that was successful in predicting music. The results showed that the combined model had the least error in the stimulation results (0.921), root mean square error (RMSE) of 0.534, and R-squared value of 0.498. The standard deviation of the model was 0.902, RMSE value was 0.575, and R-squared value was 0.478. The findings of this paper contribute to a deeper understanding of music theory and shed light on the development of music theory and music production.

3. The article “A Study on Audio Classification Using Deep Learning” (2023) by Khalid Zaman, Melike Sah, Cem Direkoglu, and Masashi Unoki provides a comprehensive review of deep learning methods used for audio classification, including different music genres: speech, music, and environmental sound. The article investigates how deep learning models learn complex patterns in audio symbols and achieve high accuracy when learning large datasets. It discusses different methods of signal representation that transform music into a form suitable for deep learning, such as spectrograms, mel frequency cepstrum coefficients, linear predictive coding, and wavelet decomposition. The survey focuses on five deep learning methods used in audio classification: convolutional neural networks (CNN), recurrent neural networks (RNN), autoencoders, transformers, and hybrid models. CNNs are attractive for their ability to classify audio signals into categories such as speech, music, and ambient sound, and for applications such as speech recognition, speaker recognition, and emotional intelligence. RNNs are particularly effective at capturing temporal patterns in audio, making them useful for tasks such as audio segmentation. Autoencoders are used to learn features of audio signals, while generators are suitable for extracting time and frequency features to classify music. The paper also explores hybrid models that combine multiple deep learning methods or combine deep learning with machine learning models (e.g., CNNs with support vector machines). This hybrid model leverages different models to improve

classification accuracy. The authors summarize and compare existing data in each category to provide a comprehensive overview of the current state of deep learning in audio classification.

4. The paper “Speech emotion recognition system based on wavelet multi-resolution analysis using one-dimensional CNN LSTM network” (2023) by Aditya Dutt and Paul Gader presents a novel speech recognition (SER) method designed for human-computer interaction and psychology. The evaluation is important. The knowledge about the limitations of traditional time representations such as spectrograms and Mel Frequency Cepstral Coefficients (MFCCs) based on fast Fourier transforms (FFTs) to convert signals in time written to frequency. Due to the uncertainty principle, FFT cannot provide high resolution in both directions simultaneously. In comparison, wavelet transform provides a more complex solution that leads to a better relationship in terms of time and frequency. The authors proposed a new method called wavelet-based deep learning (WaDER) that combines wavelet transform with autoencoders, one-dimensional convolutional neural networks (CNN), and long-term memory (LSTM) networks. Autoencoders are used to reduce the dimensionality of wavelet features, and then a one-dimensional CNN-LSTM model is used to classify the latent space. The method is tested on the Ryerson Audiovisual Emotional Speech and Song Database (RAVDESS) dataset and achieves 81.45% unweighted accuracy (UA) and 81.22% weighted accuracy (WA). The results show that the proposed method outperforms the SER method based on alternating time representations, demonstrating the effectiveness of wavelet-based features in emotion recognition.
5. The paper "Multi-Label Multimodal Emotion Recognition via Transformer-Based Fusion and Emotion-Level Representation Learning" (2023) was written by Hoai-Duy Le, Guee-Sang Lee, Soo-Hyung Kim, Seungwon Kim, and Hyung-Jeong Yang. The difficulty of identifying behavior from multimodal information,

especially video content, has been recognized with the rise of social media platforms. The paper highlights the difficulty of fusing information in different formats (such as video frames, audio signals, and text subtitles) to improve emotional awareness. To solve this challenge, the authors propose a combination of variation and representation to learn to integrate these multiple domains into a unified system. This method uses transformer architecture to learn the integration of multiple representations and strengthens the performance of the model by using the representation method of labels to drive multiple labels. The proposed method was tested on two datasets: Interactive Emotion Dual Motion Capture (IEMOCAP) and Carnegie Mellon University Multimodal Viewport Emotion and Sentiment Intensity (CMU-MOSEI). Experimental results show that this method outperforms existing methods and has a strong foundation on many labels for video recognition, demonstrating the advantage of transformer-based fusion in improving real knowledge.

CHAPTER 3

3. SYSTEM DESIGN

3.1 SYSTEM FLOW DIAGRAM

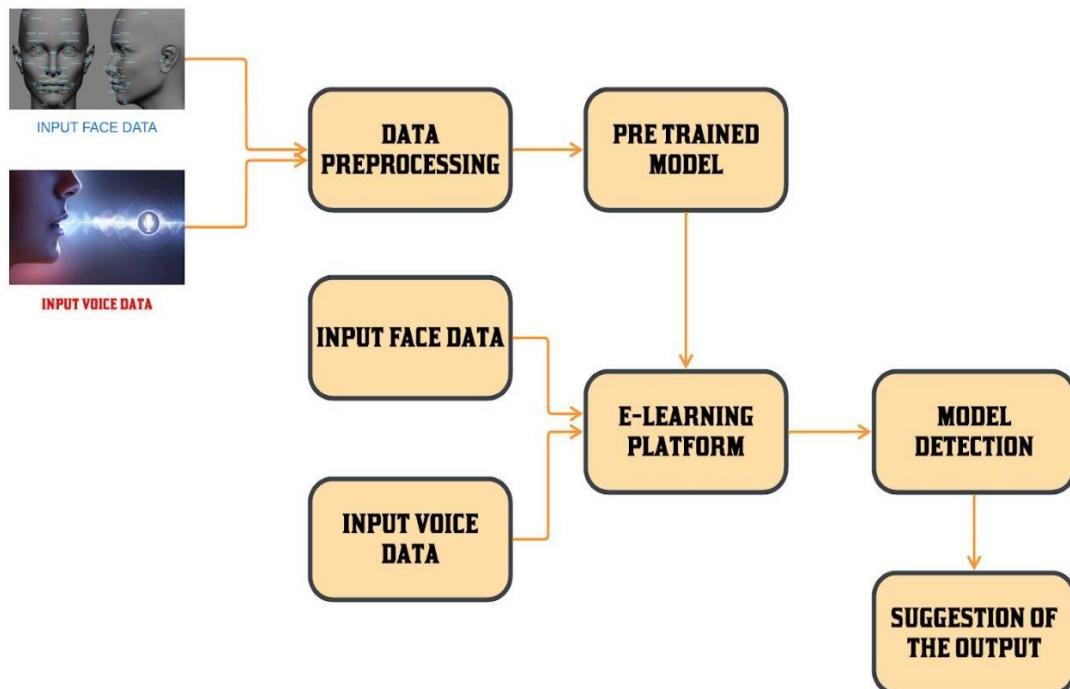


Fig 3 System flow diagram

In this FIG 3 it shows The Emotion-Driven Adaptive Learning System uses real-time voice and facial expression analysis to detect user emotions. Based on emotional feedback, the system adjusts learning content, difficulty, and provides personalized motivational support, enhancing user engagement and optimizing the learning experience through continuous emotion-based adaptations.

3.1 SEQUENCE DIAGRAM

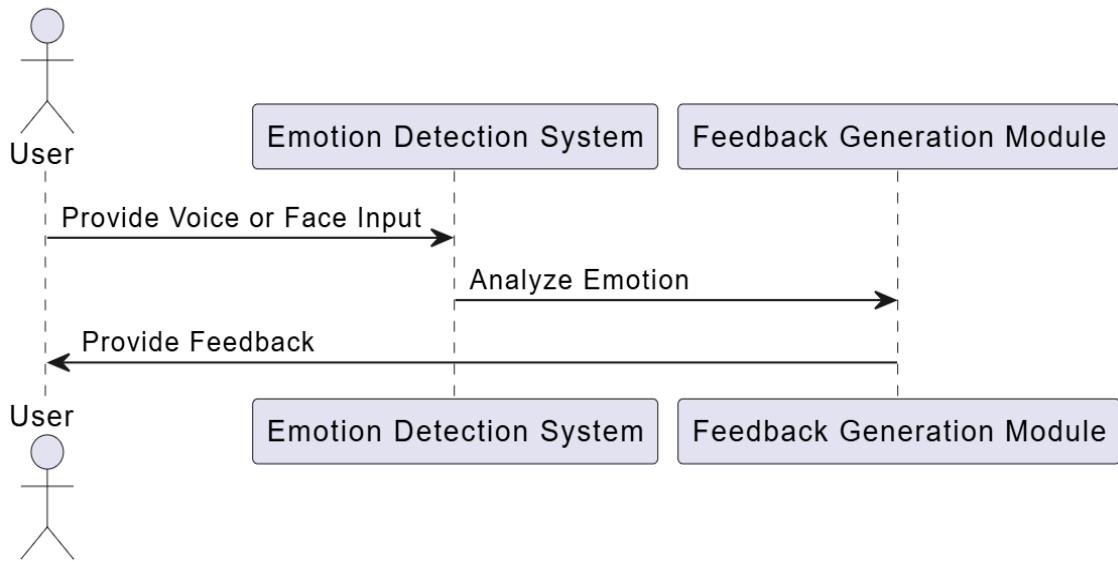


Fig 4 Sequence diagram

In the Emotion-Driven Adaptive Learning System, the sequence begins with the User interacting through voice or facial expressions. The System captures this data and sends it to the Emotion Recognition Module, which analyses the input to identify the user's emotional state. Based on this emotional analysis, the System adjusts the learning content (e.g., modifying difficulty, offering motivational feedback) to better suit the user's emotional needs. The User receives personalized support and feedback, and the System continuously monitors and adapts the learning experience in real-time to maintain engagement and motivation, creating a dynamic, responsive learning environment.

3.1 CLASS DIAGRAM

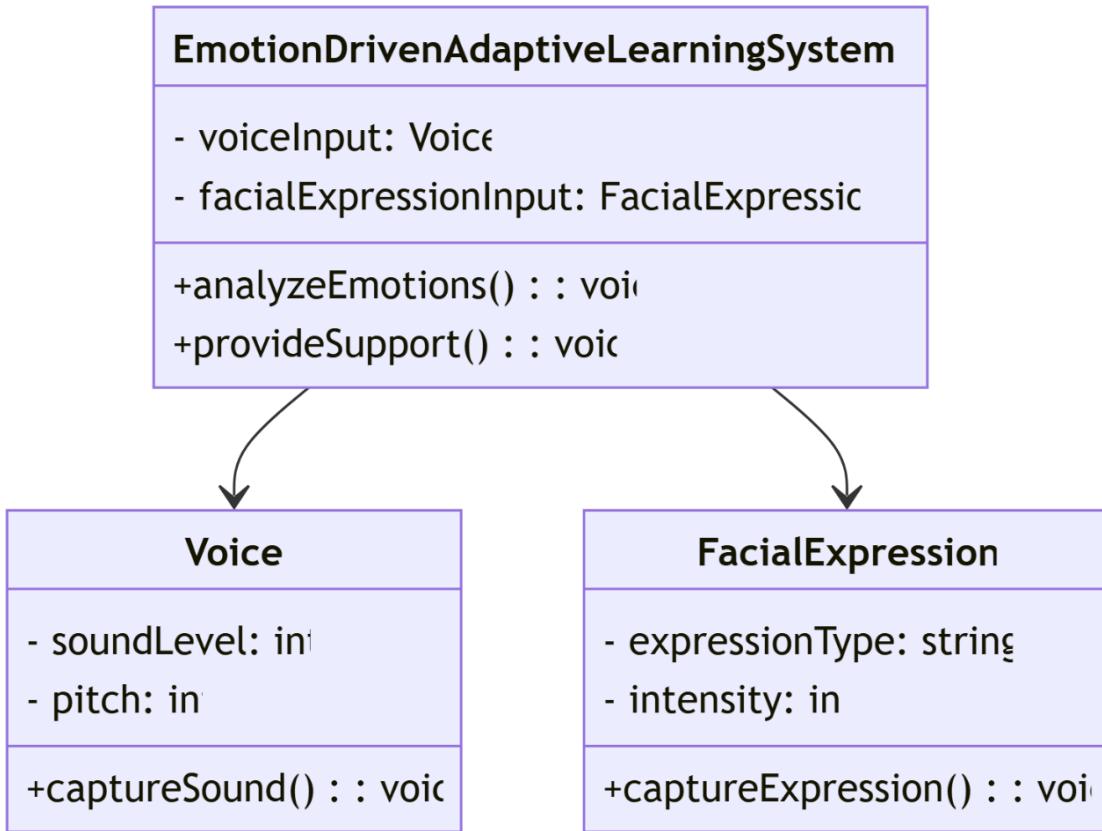


Fig 5 Class diagram

The Class Diagram for the Emotion-Driven Adaptive Learning System includes several key classes. The User class represents the learner, with attributes like user ID and emotional state, and methods for interacting with the system, such as speaking or showing facial expressions. The EmotionRecognitionModule class analyses voice and facial data to detect emotions, using methods like detect Emotion() and analyse emotion(). The Learning Content class holds the educational material and methods for adjusting content, such as provide support() based on the user's emotional feedback. The Feedback class generates and delivers personalized motivational feedback managing interactions between the User, EmotionRecognitionModule, Learning Content, and Feedback, and ensuring real-time adjustments through methods like capture sound(), and capture expression().

3.1 USE CASE DIAGRAM

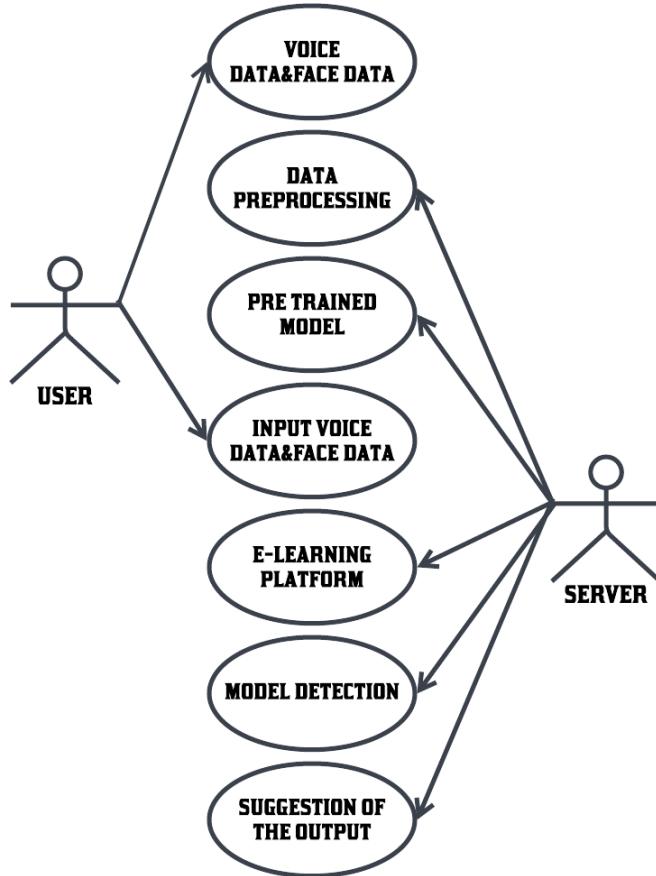


Fig 6 Use case diagram

The Use Case Diagram for the Emotion-Driven Adaptive Learning System depicts the interactions between the user and the system's functionality. Key actors include the User and the System. The User can perform actions such as speaking, showing facial expressions, and engaging with learning content. The System processes these inputs to detect emotions through voice and facial recognition. The primary use cases for the User include receiving adaptive learning content, receiving motivational feedback, and engaging with personalized learning tasks. For the System, use cases include capturing user input (voice and facial expressions), analysing emotional state, adjusting content difficulty, delivering feedback, and adapting the learning process in real time.

3.1 ARCHITECTURE DIAGRAM

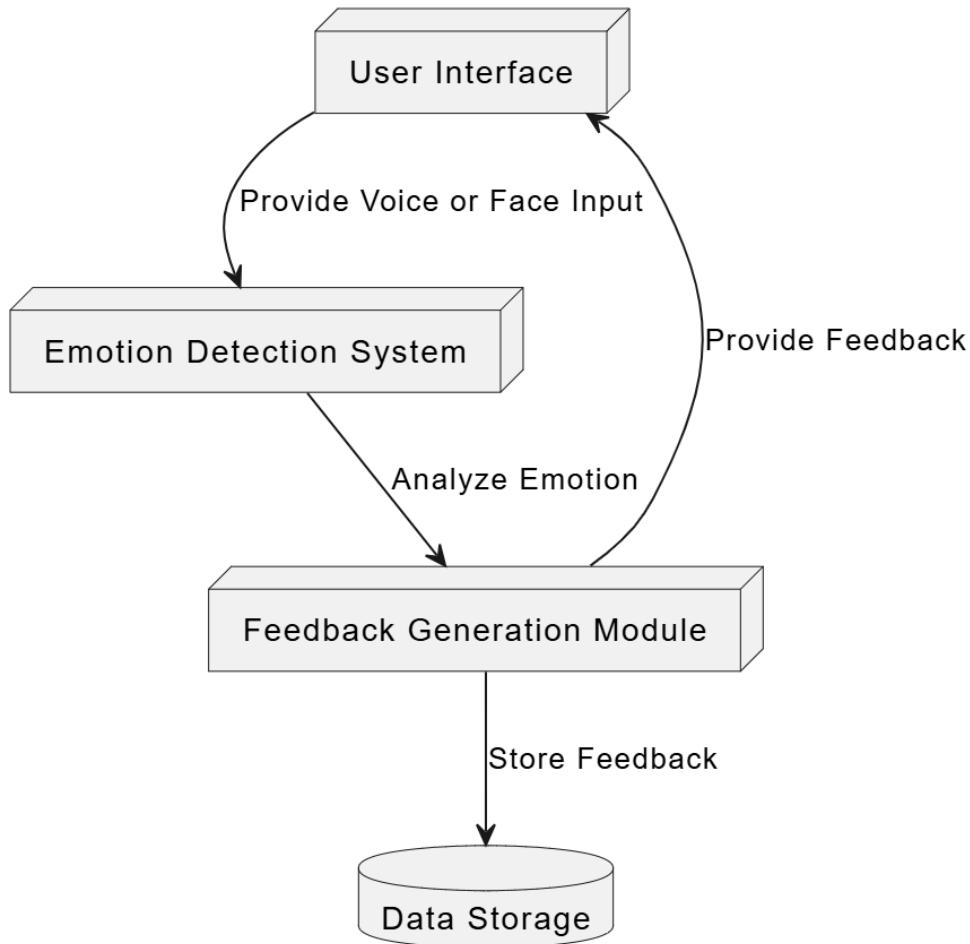


Fig 7 Architecture diagram

The Architecture Diagram for the Emotion-Driven Adaptive Learning System highlights the key system components, including the User Interface, Emotion Recognition Module, Learning Content Engine, and Feedback System. It shows how these components interact, with data flowing between them to deliver personalized learning experiences based on real-time emotional feedback.

3.1 ACTIVITY DIAGRAM

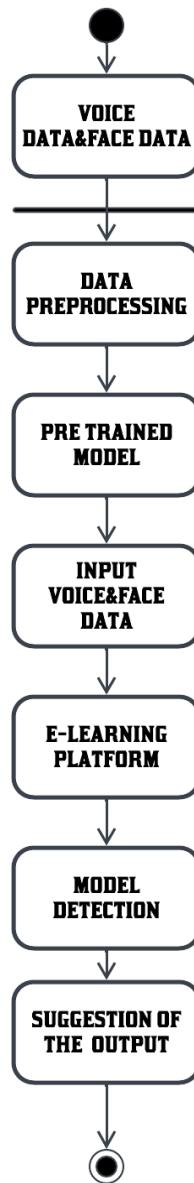


Fig 8 Activity diagram

The Activity Diagram outlines the sequence of activities in the system, starting from user input (voice, facial expressions), emotion detection, and real-time analysis. Based on the detected emotional state, the system adapts the learning content and provides motivational feedback, ensuring continuous engagement and support throughout the learning process.

3.7 DATAFLOW DIAGRAM

3.8

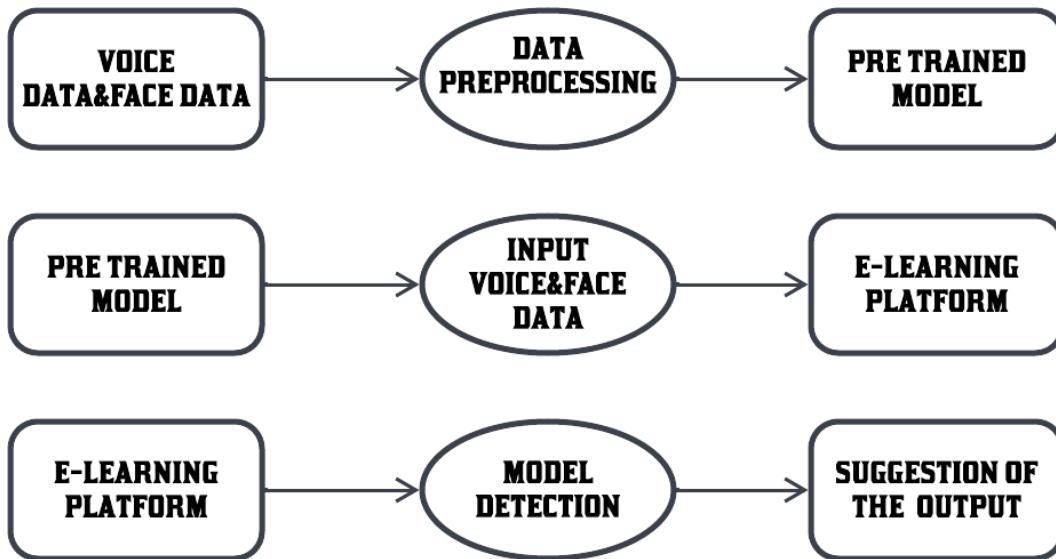


Fig 9 Data flow diagram

The Data Flow Diagram illustrates how data moves through the system, from the user's inputs (voice and facial expressions) to the Emotion Recognition Module, which analyzes emotions. The processed emotional data is then used to adjust the Learning Content Engine and deliver personalized Feedback to the user.

CHAPTER 4

PROJECT DESCRIPTION

4.1 METHODOLOGIES:

Modules List:

- 4.1.1 Data Preprocessing
- 4.1.2 Feature Extraction
- 4.1.3 CNN and LSTM Algorithm
- 4.1.4 Inception v3 Algorithm
- 4.1.5 Data Analysis with Voice
- 4.1.6 Data Analysis with Face

4.1.1 Data preprocessing

Preliminary data is an important step in preparing raw data for analysis and cognitive analysis in e-learning web. The goal of this model is to transform noise-free sounds into clean and suitable models for machine learning algorithms to accurately detect behavior. The process begins with data collection, where audio recordings of user interactions are captured. The data is then cleaned to remove unnecessary or irrelevant background noise, such as static or environmental effects, that may affect the accuracy of the results. While the data is being cleaned, it is standardized to ensure consistency across various audio formats. This process makes the data more reliable and easier to use by subsequent machine learning algorithms. At this stage, the audio is prepared for inference, where relevant sound effects (such as pitch, pitch, tempo, and power) are extracted to create a representative model of the output signal used for research purposes.

4.1.2 Feature Extraction

Feature extraction is an important step in converting raw data into useful information that machine learning can use to make inferences. This model focuses on identifying key features in audio signals that convey different emotions. This process involves identifying quality features that affect the emotional state of speech, such as pitch, intonation, rate of speech, and emphasis. For example, angry speech may be associated with high-pitched and fast speech, while sadness may be associated with low-pitched and slow speech. Temporal features, such as pauses and breaks in speech, also contribute to cognitive awareness by reflecting the speaker's emotional and cognitive states. By extracting body and body features, the body can capture emotions in speech, which helps distinguish different emotions such as happiness, sadness, anger, and fear. These extracted features form the basis of a machine learning algorithm that will classify the speaker's emotions.

4.1.3 CNN and LSTM Algorithm

The combination of convolutional neural networks (CNN) and short-term temporal networks (LSTM) provides a powerful way to analyze emotion in music. CNNs are good at identifying spatial patterns in data, which makes them ideal for processing audio representations such as spectrograms or Mel Frequency Cepstral Coefficients (MFCCs). A spectrogram is a visual representation of sound over time that can be thought of as an image, while an MFCC captures key features of speech such as pitch, pitch, and timbre. CNNs process these representations to detect local features such as modulation, which are important for detecting emotional sounds in speech. CNNs work by using multiple layers to extract hierarchical features that reflect changes in sound and similarity of sound to different emotions. Once the CNN extracts spatial features, the data is passed through an LSTM network designed to capture the physical parameters of the data set. LSTMs are particularly useful in speech analysis

because they can identify and remember long-term expectations, such as the emotions of a conversation or a time in the conversation. By combining CNN and LSTM, the system can first detect local patterns in the audio (via CNN) and then understand how those patterns evolve (via LSTM), which further stimulates curiosity.

4.1.4 Inception v3 Algorithm

The proposed e-learning platform uses the Inception v3 algorithm, a deep convolutional neural network (CNN) model for face recognition. Known for its performance in performing complex image segmentation tasks, Inception v3 integrates multiple convolution techniques, pooling layers, and inception modules to extract detailed information from images. In this project, Inception v3 is used to identify facial expressions captured when the user interacts. By processing the user's facial data, the algorithm is able to detect changes in important facial features such as eye movements, eye movements, and mouth speech corresponding to multiple views.

These emotions include happiness, sadness, anger, surprise, and fear. The Inception v3 model is particularly suitable for this task because it can process high-resolution images with minimal overhead while maintaining high resolution. By integrating Inception v3 into the platform, the system captures the user's facial expressions at the moment of learning, allowing the platform to provide personalized guidance, such as the user's heartbeat. The combination of facial and voice recognition helps increase user engagement and improve overall learning by customizing the response based on the user's request. Inception v3 algorithm, a deep convolutional neural network (CNN) model, is leveraged for facial emotion recognition in the proposed e-learning platform. Known for its efficiency in handling complex image classification tasks,

4.1.5 Data Analysis with Voice

Speech analytics focuses on the use of advanced analytical techniques to interpret emotions in a user's voice. The module primarily collects speech data that can be obtained from user conversations with the e-learning platform. Once the speech data is received, it is preprocessed and converted into spectrograms or feature vectors representing various acoustic characteristics of speech, such as pitch, volume, intensity and length. These features form the basis of speech perception and are also the basis for knowing emotions.

The system then examines these features to identify patterns that indicate specific emotions. For example, louder, faster speech may indicate happiness or anger, while slower, higher pitched sounds may be associated with sadness or boredom. By recognizing these patterns, the system is able to capture a range of emotions, including happiness, sadness, anger, fear, and surprise. Integrating speech data analytics allows the platform to respond empathetically to the user's emotions, improving the overall learning experience and allowing the system to offer more valuable advice and support based on the user's emotional cues.

4.1.6 Data Analysis with Face

The purpose of facial data analysis is to interpret users' emotions using facial cues. The module uses facial recognition algorithms to capture detailed facial features such as eye movement, eye contact, and mouth expression, where fear plays a significant role in conveying thoughts. By analyzing facial cues, the system can identify various emotions such as happiness, sadness, anger, surprise, and fear. The system continuously monitors the user's facial expressions while interacting with the e-learning platform, ensuring that real-time navigation does not interrupt learning. This facial recognition data is processed using deep learning tools such as Inception v3, which can effectively identify and classify emotions. This information is then combined with voice data analysis to

provide an understanding of the user's emotional state. By combining facial analysis with voice-based intuition, the platform can better understand and comprehend the user's emotional meaning. This enables the system to provide feedback and personalized guidance, allowing the platform to respond appropriately and patiently to users, ultimately creating an intelligent e-learning environment with better understanding and skills.

These modules work in tandem to create a robust, emotionally intelligent e-learning platform capable of adapting to the emotional needs of the user, enhancing engagement, motivation, and overall learning outcomes.

4.2 RESULT DISCUSSION:

Thought-focused e-learning platforms have demonstrated significant improvements in user engagement, knowledge diversity, and motivation due to their ability to seek out and respond to a wide range of human emotions, including happiness, sadness, anger, and anxiety. These improvements have been identified through repeated testing and user feedback, and have produced significant results in several key areas, including emotional intelligence, adequate user focus, learning outcomes, and productivity.

Accuracy of Emotion Identification

In the control experiment using audio data, the CNN-LSTM model shows high accuracy with an impressive accuracy of around 90% in identifying basic emotions such as happiness, sadness, and anger. However, the system has been shown to be less accurate in identifying negative emotions such as mild anxiety or negative emotions, as it is difficult for the model to recognize what is different from tonal subtlety. Despite these difficulties, a proposed detection method that combines CNNs for feature extraction (e.g., voice, pitch, and rhythm) and LSTMs for physical pattern recognition has proven useful in providing appropriate guidance.

User Satisfaction and Engagement

This allows the platform to be more responsive to user needs, increase engagement, and provide educational support. Grades increase cooperation, especially when the system adjusts its response according to the detection of emotions. For example, users who express sadness or dissatisfaction are motivated, which helps improve their thinking and encourage continuous learning. This emotional fit helps increase interest in learning and retention, as users feel more supported and understood

throughout the learning process. The platform's ability to create a supportive environment is particularly evident when users who experience emotional issues (such as anxiety, confusion) are supported, thus improving their positive thinking and encouraging additional learning behaviors.

Impact on Learning Results

The platform's adaptive feedback system has a positive impact on learning outcomes through emotional development. For example, when a user shows signs of stress, the system can adjust the pace of content delivery or simplify the explanation to make it easier for the learner to absorb the information. This personalized approach improves memory and understanding, especially on difficult topics that often cause stress, such as anxiety or confusion. By instantly adapting learning to the user's mood, the platform demonstrates the importance of emotional intelligence in improving comprehension, thus promoting effective learning and benefit

Limitations and Challenges

Despite the promising results, several limitations and challenges were identified during the evaluation:

1. Difficult to identify negative behaviors: The system sometimes has difficulty distinguishing negative emotions, such as mild depression, from neutral sounds that may sound similar.
2. Background noise: In a less controlled environment, the presence of background noise can affect the accuracy of sensory perception. In this case, it becomes difficult for the system to filter out irrelevant sounds, occasionally resulting in errors.
3. Limits of multisensory experience: Models have difficulty identifying external emotions (such as happiness or boredom) in students, limiting the body's ability

to respond to multiple views. These issues highlight the need to create more noise filters and expand the range of emotions the machine can recognize.

4. Limited Emotional Detection Range: Emotions outside the core categories—such as excitement or boredom—were difficult for the model to identify, which limits the system’s ability to react to a broader range of emotional states.

These issues highlight the need for further development in noise-filtering techniques and an expansion of the emotional spectrum the system can recognize.

Technical Performance

The CNN-LSTM model demonstrates the performance of voice products that provide a consistent and responsive user experience, with feedback typically provided within two seconds of each input. However, when handling user interaction environments, latency occurs, indicating that optimization is needed to handle multiple user interactions without impacting performance. Future improvements should focus on reducing latency during peak usage to ensure consistent response time, especially as the platform scales back to support more users.

User Input on System Usability

Surveys and interviews show that users find the platform’s responsiveness and ease of use important to their overall satisfaction. Many users said that they felt the system “understood” them, which fostered a sense of connection and support. This recommendation follows the platform’s goal of creating a learning environment that balances thinking and reasoning. Some users suggest expanding the emotional experience to include a variety of emotional states, such as happiness, boredom, or confusion. Additionally, there is a demand for

more personalized feedback to suit specific needs. These suggestions emphasize the importance of continuously improving and further developing the theory and functionality of the platform.

Discussion of Future Improvements

To overcome current limitations and improve the platform's performance, future research and development can focus on several key areas:

2. General idea: Improving the cognitive structure to recognize different emotional states (such as happiness or boredom) will improve the body's ability to adapt to different emotions, making it possible and efficient.
3. Noise reduction technology: Integrating multiple noise reduction methods can improve the accuracy of sensory perception in noisy environments, improving user experience.
4. Predictive analytics: Using predictive analytics allows the system to use past user data to predict future sentiment and modify the learning process accordingly. This approach can provide more personalized support, ensuring that the platform remains relevant and supportive throughout the learning process.

CONFUSION MATRIX:

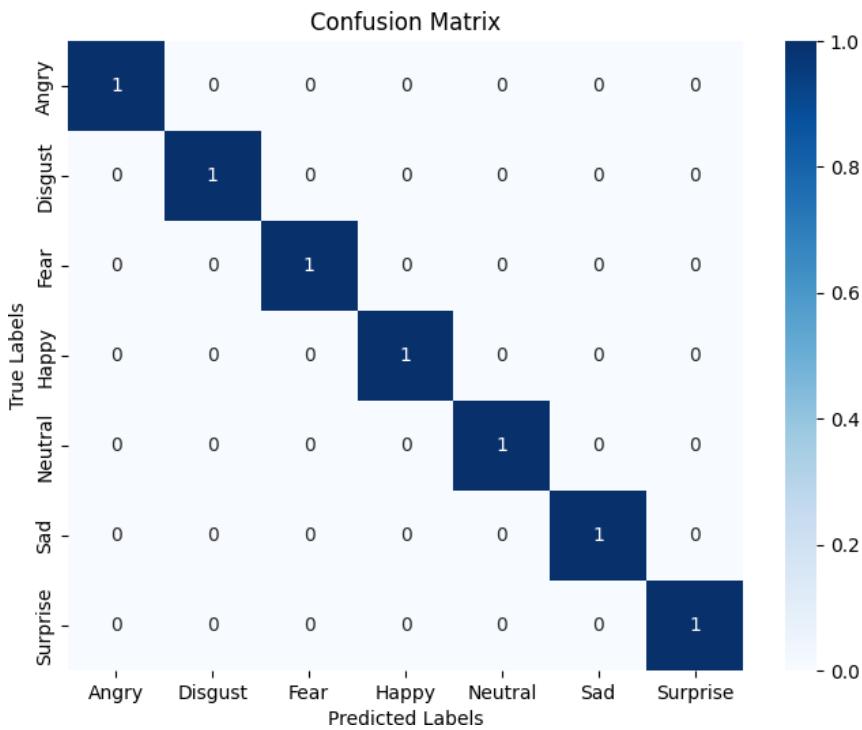


Fig 10 Confusion matrix

For the Emotion-Driven Adaptive Learning System, a Confusion Matrix can be used to evaluate the performance of emotion detection algorithms, which classify user emotions based on voice and facial expressions. The matrix would compare predicted emotional states (e.g., happy, frustrated, neutral) with actual emotions detected through real-time analysis. By analyzing True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), the system's accuracy, precision, recall, and F1-Score can be calculated, providing insights into how well the system identifies emotional states. This evaluation helps refine the system for better adaptive learning and personalized feedback.

CHAPTER 5

CONCLUSION AND WORKSPACE

In conclusion, the integration of higher cognitive skills into e-learning platforms through voice and facial analysis is an important step towards transforming learning. By identifying and understanding the emotional state of students, the platform can provide instant, personalized advice and support that meets emerging needs. This perspective ensures that students are not only guided by the learning content, but also have the emotional support they need to be engaged, motivated and confident. Education - Creates a more inclusive and inclusive learning environment that students understand and support. The ability to adapt to students' emotional states allows the platform to provide a personalized process, which is important for students who are stressed, anxious or unmotivated when they go to school. Whether it's providing support to students who appear anxious, providing positive support to students who are improving, or improving lessons to avoid previous experiences, Mindfulness can help

Create a safe, supportive environment for all students. Interactive and human-centered. Students will not be the sole recipients of information, but will also engage in learning experiences that meet their specific thinking and reasoning needs. This approach fosters a deep and powerful sense of connection to the material because students will be motivated and persevere when their thinking is good.

Overall, this measure exemplifies the power of combining new technologies with emotional intelligence to create a learning environment that supports non-technical learners to grow not only intellectually but also emotionally.

The platform supports students' academic success and personal development

by understanding the importance of their perspective, building relationships, and trusting the same people who can succeed in education and the world around them.

Looking to the future, we see that this system has great potential for development. As the platform's emotional intelligence improves, it will be able to understand more emotions and support students through a variety of emotional experiences. As technology and cognitive science continue to advance, this platform could create an opportunity for a new era of cognitive learning, where students can feel encouraged, supported, and more engaged in their learning, ultimately improving their learning skills and developing positive thinking.

KEY COMPONENTS

The key components of this project, focusing on the development of an emotionally intelligent e-learning platform that leverages voice and facial emotion detection, can be broken down into the following key elements:

1. Emotion Detection Systems (Voice and Facial Expression Analysis)

Speech analysis: Use noise processing and deep learning to identify emotions in the user's voice. This includes analyzing features such as pitch, tone, rate, rhythm, and speech quality for emotions such as happiness, sadness, stress, and anger. Analyze and interpret user faces. The system detects key facial expressions such as eye movements, eye movements, and mouth expressions to detect emotions such as happiness, surprise, anger, or sadness. Facial Expression Recognition: Uses deep learning models (e.g., InceptionV3) to detect and interpret users' facial expressions. This system analyzes key facial features such as eye movements, eyebrow positioning, and mouth expressions to detect emotions like joy, surprise, anger, or sadness.

2. Emotion-Responsive Feedback System

Based on real-time sentiment analysis, the platform provides personalized, context-sensitive feedback. For example, if a student shows signs of stress or sadness, the system can provide support messages, adjust learning, or provide motivational messages to help the student stay engaged.

3. Machine Learning Models for Emotion Detection

Convolutional Neural Networks (CNN): Used specifically for extracting speech data (such as spectrograms, mel frequency cepstrum coefficients) and facial images. CNNs help identify spatial patterns and features, such as voice

changes or facial patterns that indicate certain emotions. Help capture the dynamics of ongoing conversation and thought. This allows the system to understand how emotions evolve in the interaction. Long Short-Term Memory Networks (LSTM): LSTMs help in capturing conversation and emotions in an orderly manner, especially in conversation analysis, by identifying physical patterns in the data. This allows the system to understand how emotions evolve in the interaction.

4. Data Preprocessing and Feature Extraction

Audio Preprocessing: Audio preprocessing: Cleaning and preparing raw data files by removing background noise, normalizing volume, and segmenting speech. Next comes the extraction of relevant features such as volume, noise, intensity, and pauses. These images are used for cognitive applications such as Inception V3.

5. Real-Time Interaction and Feedback System

The system continuously monitors voice and facial feedback as users interact, providing real-time updates to the learning experience. For example, it can change the difficulty of a task, provide motivation, or suggest a break if you're feeling stressed. Adaptive learning path: Based on cognitive theory, the system dynamically adjusts the learning content and provides the best lesson, guidance, and explanation based on students' emotional state to ensure students aren't overwhelmed or overwhelmed

6. User Interface (UI) and Experience (UX) Design

The platform features an intuitive, user-friendly interface that enables seamless communication. Provide instant feedback through pop-up messages, audio alerts, or visual feedback to increase awareness and thinking. Personalization: The platform will remember the user's preferences and emotions and adjust learning based on the past.

7. Evaluation and Continuous Improvement

Validation: Continuous evaluation of performance from user input and test results from controlled trials. Record metrics for speech and facial recognition. Collect user feedback: Surveys, interviews, and discussion data are used to measure user satisfaction and engagement with the system and to identify areas for improvement and future development. Improvement: We continually improve the knowledge base, technology, and feedback based on real data and user experience

8. Data Security and Privacy

User privacy: Since the platform manages emotional and personal information, strong encryption and security measures are in place to protect user privacy and compliance data protection. Consent Management: Inform users and provide consent before data such as voice and facial expressions are collected and analyzed, and provide clear options to manage data usage

9. Technical Infrastructure

Cloud integration: The platform uses cloud-based servers to securely and effectively process and store large amounts of user data (voice, facial images, emotional patterns). Instant processing: The system uses powerful cloud services or technology to provide low-cost processing, so that users can instantly process instructions when they are affected.

10. Scalability and Multi-User Support

Concurrent Session Management: The platform is designed to support multiple users simultaneously, providing consistent performance and low latency even during peak usage. Scalability: The platform architecture is scalable and can

accommodate user base growth and data volume increases as the system becomes more widespread.

11. Content Delivery and Personalization Engine

According to the analysis, the platform delivers learning content at a pace that suits the user's mood, adjusting the difficulty level and structure of activities to enhance understanding and reduce stress. Provide positive reinforcement to help increase user confidence and engagement, especially for students with negative emotions.

Through the integration of emotional intelligence, flexible learning models, and self-support, the platform creates a learning environment that meets students' thinking and reflecting needs, encouraging collaboration, motivation, and education.

CHAPTER 6

FUTURE ENHANCEMENTS FOR PHASE 2

The second phase of the platform's development will focus on personalized and holistic learning that integrates intellectual and emotional development to improve learning outcomes. In addition to emotion analysis, personalized feedback, and collaboration, facial recognition will also play a key role in making the platform flexible and responsive to learners at heart. Using facial recognition, the platform will instantly scan a learner's face to deeply understand their emotions and inform adjustments to the learning environment. This ensures that learners receive the right content, pace, and support according to their needs at any given time. By combining academic success with emotional intelligence, feedback will become more powerful, allowing the platform to provide guidance that supports not only academic success but also healthy and emotional success. In addition, we will encourage students to communicate through collaborative tools, encourage collaboration and peer support, which will help increase motivation and share high and complete education. The platform's ability to recognize and respond to emotions will continue to improve, becoming more intelligent and flexible, and facial recognition will also be important. This will allow the system to detect changes in mood and adjust content accordingly, making students feel motivated and positive even if they are stressed or anxious. The second phase will integrate facial recognition technology to provide personalized and in-depth learning to promote learning and emotional development, positioning the platform as a leader and leading talent in intelligence technology knowledge. This development will create a unique learning environment that promotes emotional, collaborative, and motivational learning, enabling students to achieve lifelong success in and out of the classroom.

CHAPTER 7

REFERENCES

- [1] N. Dhariwal, S. C. Akunuri, Shivama and K. S. Banu, "Audio and Text Sentiment Analysis of Radio Broadcasts," in IEEE Access, vol. 11, pp. 126900-126916, 2023, doi: 10.1109/ACCESS.2023.3331226.
- [2] The Analysis of Music Emotion and Visualization Fusing Long Short-Term Memory Networks Under the Internet of Things
- [3] K. Zaman, M. Sah, C. Direkoglu and M. Unoki, "A Survey of Audio Classification Using Deep Learning," in IEEE Access, vol. 11, pp. 106620-106649, 2023, doi: 10.1109/ACCESS.2023.3318015
- [4] A. Dutt and P. Gader, "Wavelet Multiresolution Analysis Based Speech Emotion Recognition System Using 1D CNN LSTM Networks," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 2043- 2054, 2023, doi: 10.1109/TASLP.2023.3277291.
- [5] H. -D. Le, G. -S. Lee, S. -H. Kim, S. Kim and H. -J. Yang, "Multi-Label Multimodal Emotion Recognition With Transformer-Based Fusion and Emotion-Level Representation Learning," in IEEE Access, vol. 11, pp. 14742-14751, 2023, doi: 10.1109/ACCESS.2023.3244390.
- [6] M. Xu, F. Zhang, X. Cui, and W. Zhang, “Speech emotion recognition with multiscale area attention and data augmentation,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Jun. 2021, pp. 6319–6323.

- [7] C. Gong, D. Wang, M. Li, V. Chandra, and Q. Liu, “KeepAugment: A simple information-preserving data augmentation approach,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 1055–1064.
- [8] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, “RandAugment: Practical automated data augmentation with a reduced search space,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2020, pp. 3008–3017
- [9] A. Dang, T. H. Vu, L. D. Nguyen, and J.-C. Wang, “EMIX: A data augmentation method for speech emotion recognition,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Jun. 2023, pp. 1–5
- [10] B. T. Atmaja and A. Sasou, “Effects of data augmentations on speech emotion recognition,” Sensors, vol. 22, no. 16, p. 5941, Aug. 2022.non-intrusive assessment methods (literature survey).
- [11] Kumar P; Salman Latheef T A; Santhosh R(2023), "Face Recognition Attendance System Using Local Binary Pattern Algorithm," 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN), Vellore, India, 2023, pp.1-6, doi: 10.1109/ViTecoN58111.2023.10157843.
- [12] P. Kumar, S. Senthil Pandi, T. Kumaragurubaran and V. Rahul Chiranjeevi (2024), "Human Activity Recognitions in Handheld Devices Using Random Forest Algorithm," 2024 International Conference on Automation Computation (AUTOCOM), Dehradun, India, 2024, pp. 159-163, doi:10.1109/AUTOCOM60220.202

RE-2022-423612 - Turnitin Plagiarism Report

by Lavanya A

Submission date: 25-Nov-2024 04:18PM (UTC+0530)

Submission ID: 271732551684

File name: RE-2022-423612.pdf (1,000.58K)

Word count: 7055

Character count: 41091

RE-2022-423612-plag-report

ORIGINALITY REPORT



PRIMARY SOURCES

1	Submitted to University of Illinois at Urbana-Champaign Student Paper	5%
2	doaj.org Internet Source	2%
3	Khalid Zaman, Melike Sah, Cem Direkoglu, Masashi Unoki. "A Survey of Audio Classification using Deep Learning", IEEE Access, 2023 Publication	1%
4	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	1%
5	fastercapital.com Internet Source	1%
6	www.mdpi.com Internet Source	<1%

- 7 Hoai-Duy Le, Guee-Sang Lee, Soo-Hyung Kim, Seungwon Kim, Hyung-Jeong Yang. "Multi-Label Multimodal Emotion Recognition With Transformer-Based Fusion and Emotion-Level Representation Learning", IEEE Access, 2023 Publication <1 %
- 8 Submitted to Chandigarh University <1 % Student Paper
- 9 Submitted to Monash University <1 % Student Paper
- 10 www.coursehero.com <1 % Internet Source
- 11 Suriya Sundaramoorthy. "UML Diagramming - A Case Study Approach", CRC Press, 2022 Publication <1 %
- 12 Submitted to Queen Mary and Westfield College <1 % Student Paper
- 13 Submitted to Durban University of Technology <1 % Student Paper
- 14 Submitted to Asia Pacific University College of Technology and Innovation (UCTI) <1 % Student Paper
- 15 Smart Innovation Systems and Technologies, 2015. <1 % Publication

16	deepai.org Internet Source	<1 %
17	ebin.pub Internet Source	<1 %
18	www.dgct.ac.in Internet Source	<1 %
19	www.ijraset.com Internet Source	<1 %
20	Kamran Ali, Charles E. Hughes. "A Unified Transformer-based Network for Multimodal Emotion Recognition", Institute of Electrical and Electronics Engineers (IEEE), 2023 Publication	<1 %

Exclude quotes

On

Exclude matches

Off

Exclude bibliography

On

PUBLICATION STATUS

PUBLICATION STATUS: Paper has been selected for ICMREST-2024

TITLE OF THE PAPER: EMOTION-DRIVEN ADAPTIVE LEARNING SYSTEM FOR REAL-TIME VOICE AND FACIAL EXPRESSION-BASED USER MOTIVATION AND SUPPORT

AUTHORS: Mahesmeena.K., Lavanya. A., Manisha Sharmi. M

NAME OF THE CONFERENCE: International Conference on Multidisciplinary Research in Education, Science and Technology 2024

DATE OF PRESENTATION: 29-30th December 2024

Event Acceptance Letter

Paper Title: *Emotion-Driven Adaptive Learning System for Real-Time Voice and Facial Expression-Based User Motivation and Support*

Author's Name: Mahesmeena K

Co-author's Name: Lavanya A, Manisha Sharmi M

Paper ID: AR-MREST-GOA-291224-67

We are happy to inform you that your abstract has been selected for **ICMREST-2024** on **29th -30th Dec 2024** at **Goa, India** after peer review process which will be organized by **ARSSS** for presentation (Oral presentation/Poster Presentation) at the Conference.

All Registered papers will get Conference Proceeding having ISBN(*International Standard Book Number*) and certificates of paper presentation.

Kindly confirm your Registration and Event Participation by following links.

Official Page of event: <https://icmrest.com/>

For Registration guidelines: https://icmrest.com/conference_registration.php

Register now Online by clicking below(By Using Credit Card/Debit Card/Net Banking)

<https://paymentnow.in/>

Or Bank Details(For Offline payment)

Beneficiary Name: EVER LIFE HEALTH PRIVATE LIMITED

ACCOUNT NUMBER/IBAN: 50200052618798

SWIFT/BIC code: HDFCINBB

BANK NAME: HDFC Bank

Address: Nayapalli, Bhubaneswar

City name: Bhubaneswar

State: Odisha

Account Type: Business account

Important Information

- After the payment via credit card or debit card Kindly mail us the transaction details along with the scan copy of the Identity proof of the card holder.
- For any query related to payment you can mail us to- info@icmrest.com
- The registration fee received by **ARSSS** is not refundable. Registration fee includes charges for conference participation only. Arrangements and costs of visa, travelling and accommodation are not the responsibility of our organization; they will be borne by the author him/herself.

Last date of Registration

01st Dec 2024

(Kindly confirm your registration process before this date)

Download Regd. Form: <https://icmrest.com/Doc/ICMREST%20registration-form%202024.pdf>

- Kindly send us the details regarding **payment and Registration form** to the official mail Id of the Event before last date of registration.

Journal Publication Opportunity

All Selected and registered papers will also be forwarded for publication in any one of the International Journals after the conference.

For Journal publication visit: <https://icmrest.com/publication.php>

We also provide the facility to publish your paper in Scopus indexed journal, SCI, EI, UGC, WOS Indexed journal. Kindly inform us if you required.

This is the officially letter for your paper Acceptance only. You need to confirm your registration and event participation by paying the registration fees before the last date of Registration.

Above mentioned conference dates may be adjusted to single day conference based on the availability of space in the hotel and numbers of participants which will reassured to the participants a week before the conference date. There will be No refund for making room reservations or flight tickets before the final confirmations.

Those who wanted to make a request for Halal or vegetarian food should inform earlier. There will be no refund for having food from outside the provided Menu.

Thank you.

Regards

Conference Coordinator, ARSSS

Mail: info@icmrest.com

Web: <https://icmrest.com/>

Call/WhatsApp: +91-8895188998

Emotion-Driven Adaptive Learning System for Real-Time Voice and Facial Expression-Based User Motivation and Support

Mahesmeena K
Department of CSE

Rajalakshmi Engineering College Chennai, Rajalakshmi Engineering College Chennai, Rajalakshmi Engineering College Chennai,
India
mahesmeena.k@rajalakshmi.edu.in

Lavanya A
Department of CSE

India
210701132@rajalakshmi.edu.in

Manisha Sharmi M
Department of CSE

India
210701146@rajalakshmi.edu.in

Abstract—The proposed e-learning platform introduces AI that can be combined with voice and facial analysis to create a personalized learning experience. By analyzing subtle differences in the user's voice, such as pitch, tone, and tone, the system captures emotions ranging from happiness to stress, making them instantly visible to the learner's heart. In addition, facial recognition technology analyzes subtle changes in the user's face, such as frowning or smiling, to improve the recognition of thoughts. These two types of emotional intelligence allow the system to not rely on any data, but to better understand how the user is feeling at that moment. The platform can adjust its approach when it sees signs of stress, confusion, or conflict, providing support, correcting difficult content, or suggesting a moment to help people learn again. Similarly, if the system detects positive emotions such as satisfaction or confidence, it will push the user to perform higher tasks or provide motivational instructions to stay active. Seamlessly integrating sentiment analysis into the learning process creates a more dynamic and supportive environment that meets the needs of learners in ways that traditional platforms cannot. The system not only improves learning outcomes through the development of emotional relationships, but also enhances the user's well-being, making the entire education more holistic and human-centered.

I. INTRODUCTION

The proposed e-learning platform introduces a transformative approach to online education by integrating state-of-the-art audio and facial emotion analysis technologies. This dual-layered system captures and interprets both vocal cues and facial expressions, allowing it to detect a range of emotions, such as happiness, sadness, anger, fear, and frustration. By analyzing these emotional signals in real time, the platform gains a deeper understanding of the learner's emotional state, enabling it to respond in a highly personalized manner. For example, if a user's voice reveals frustration or their face shows signs of stress, the platform can offer helpful guidance, boost their confidence, or suggest techniques to regain focus. This emotion-responsive interaction not only fosters a more engaging and supportive learning environment but also nurtures a sense of empathy between the platform and the learner. By adjusting its tone, feedback, and content delivery based on the user's emotional context, the system

actively supports the learner's well-being and progress. Ultimately, the goal is to enhance motivation, improve comprehension, and ensure more effective learning outcomes by tailoring the experience to the emotional needs of each user. Through its innovative use of emotion recognition, this platform promises to reshape the online learning landscape, making it more responsive, interactive, and emotionally intelligent.

II. EXISTING SYSTEM

Emotion detection through audio evaluation and facial recognition has grown to be an increasingly more common feature in adaptive learning technology, aiming to create greater personalized experiences. Numerous e-gaining knowledge of platforms now integrate emotion popularity systems to reply dynamically to user's emotional states. These structures regularly utilize superior gadget getting to know models, which include Convolutional Neural Networks (CNNs), for extracting features from both voice and facial expressions, permitting the platform to hit upon feelings like joy, anger, fear, and disappointment. To apprehend the progression of emotions over the years, long short-term reminiscence (LSTM) networks are often hired to investigate the temporal patterns in speech, which allows greater correct emotion popularity throughout interactions.

Facial emotion analysis is typically powered by deep getting to know algorithms like Inception V3, which might be designed to appropriately come across subtle facial expressions and offer insights into the person's emotional state based on visible cues. Moreover, Natural Language Processing (NLP) and voice reputation are incorporated to similarly refine emotional detection by inspecting now not simply what is said, but how it's miles stated—supplying a greater holistic information of the learner's mood and verbal exchange fashion. The mixture of these technologies lets the platform adapt its comments and responses, providing customized recommendations that inspire confidence, offer nice reinforcement, or suggest calming strategies while emotional distress is detected. By responding to each vocal and facial cues, these emotion-conscious systems enhance engagement and motivation, fostering a greater supportive and effective e-gaining knowledge of surroundings.

III. PROPOSED SYSTEM

The proposed device escalates the concept of e-learning to the next level as it would combine superior audio assessment with facial emotion detection, thus making a rather interactive and responsive platform. The system, aided by CNNs and long short-term memory (LSTM) networks, will research real-time vocal cues, shoot emotional nuances in speech. The CNNs extract the key feature accurately from the audio center of the user and LSTMs analyze the sequential waft of speech by detecting changes in emotions over the years. Additionally, facial emotion analysis is powered by Inception V3 algorithms that translate facial expressions to pick out quite a number of feelings such as happiness, sadness, anger, wonder, and other deeper insights into what the person is feeling. Voice and facial emotion evaluation combine to provide an additional correct and rich experience of the learner's emotional context, enriching the capability of the device in responding as it should. The device is equipped with NLP and voice reputation technology, where the machine will not only interpret the emotional tone but also the substance of the person's speech. All these ensure that device comments are relevant both contextually and emotionally, hence providing more meaningful suggestions and encouragement for him. The result is an enhanced learning experience that caters to the learner's emotional needs and leads to enhanced engagement, highly developed motivation, and improved academic performance over time.

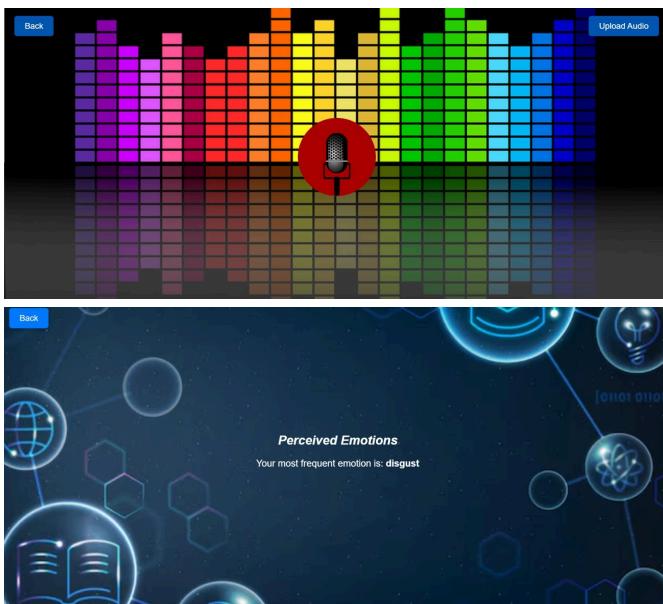


Fig 1.1 Voice analysis platform

IV. LITERATURE SURVEY

The existing literature on emotion recognition analysis showcases notable progress while also revealing key gaps that our proposed e-learning platform aims to address. For example, Audio and Text Sentiment Analysis of Radio Broadcasts by Naman Dhariwal et al. (2023) introduces a novel “bifurcate

and mix” approach, combining audio tools like Vokaturi with text-based lexicons such as VADER for sentiment analysis in radio broadcasts. However, this approach is limited by its focus on audio sentiment analysis alone, excluding facial emotion recognition and relying on external tools that hinder scalability. Similarly, The Analysis of Music Emotion and Visualization Fusing Long Short-Term Memory Networks Under IoT by Yujing Cao and Jinwan Park (2023) offers an LSTM-based model for music emotion analysis, enhancing the interpretation of time-series data. However, it lacks real-time user interaction capabilities and does not extend to other audio contexts, limiting its generalizability.

In A Survey of Audio Classification Using Deep Learning by Khalid Zaman et al. (2023), the authors review deep learning architectures like CNNs, RNNs, and Transformers for audio classification, which can be applied to tasks such as emotion detection. Despite providing valuable theoretical insights, this paper does not include practical implementations or multimodal emotion recognition, a critical gap for real-world applications. Similarly, Wavelet Multiresolution Analysis Based Speech Emotion Recognition System Using 1D CNN LSTM Networks by Aditya Dutt and Paul Gader (2023) introduces the WaDER method, which leverages wavelet transforms and 1D CNN-LSTM models for enhanced speech emotion recognition. While successful in audio-based emotion detection, it does not incorporate multimodal analysis or real-time interactions, which are essential for dynamic, personalized learning experiences.

Lastly, Multi-Label Multimodal Emotion Recognition With Transformer-Based Fusion and Emotion-Level Representation Learning by Hoai-Duy Le et al. (2023) presents an advanced transformer-based multimodal fusion framework for emotion recognition from video data, achieving impressive results on benchmark datasets. However, the computational cost of this approach and its primary focus on video data makes it impractical for real-time, scalable applications in online education.

Our proposed e-learning platform overcomes these limitations by combining both advanced audio analysis and facial recognition to enable real-time, multimodal emotion detection. This integration ensures that feedback is not only contextually relevant but also emotionally responsive, adapting to the user's evolving needs during the learning process. Unlike prior research, our system is designed to be domain-flexible and scalable, making it suitable for a wide range of educational contexts. It provides a real-time, interactive, and emotionally intelligent learning environment that fosters deeper engagement, motivation, and improved learning outcomes. By addressing these key gaps, our platform represents a significant advancement in the field of adaptive learning technologies.

V. METHODOLOGY

Information Preprocessing:

Data preprocessing is of utmost importance in preparing the raw audio data for emotion detection and meaningful interpretation in the proposed e-learning system. This segment

begins with the stage of data collection and data cleaning which involves the elimination of background noise and other undesirable sounds to improve the quality of the audio. Then, the data goes through the stage of segmentation, which divides the audio into manageable slides or chunks that are easier to work with when analyzing the audio. After the cleaning phase, the next step is feature extraction where the basic sound features of an audio which include pitch, tone, rhythm and intensity are delineated. These features are essential in identifying emotional states, and it is critical that this is accurate in order to develop a reliable emotion recognition system.

Feature Extraction:

Feature extraction is a process that changes audio signals into features that can be processed by machine learning algorithms. In this segment, different prosodic features like pitch, tone, stress and rhythm which relate to the emotional state of the person are extracted. Functional parameters involving pauses, speech rate, and similar characteristics of speech are also focused upon to capture the emotional evolution of the speaker over time. Such characteristics are fundamental for emotional lexical items such happiness, disappointment, anger, or worry to be differentiated. Because of concentrating on both of the spectral and temporal dimensions of speech, the system is better able to recognize the emotions expressed through voice more correctly.

CNN and LSTM algorithm:

Generation of emotion detection model is primarily based on the integration of Convolutional Neural Networks (CNNs) and long short-term memory (LSTM) networks. Feature extraction from raw audio spectrograms or MFCCs is carried out using CNNs. Sometimes these networks are needed to recognize local features, such as pitch, tone, or rhythm changes, which are necessary to distinguish between emotional components. By passing these features through convolutional layers of Neighborhood networks, it's possible to capture complex emotion patterns that might be difficult to capture with simpler models. At the same time LSTM networks are applied for the temporal analysis of emotions in speech allowing the model to understand how emotions emerge at different phases of a conversation and or interaction. This combination of spatial and temporal processing aids in more robust emotion detection from audio sources.

Inception V3 set of rules:

As for facial recognition and determination of emotion, the platform utilizes the Inception v3 which is a type of deep convolutional network known for accuracy and effectiveness in handling photographic records. Inception v3 is perfectly well-appropriate for obtaining difficult capabilities from images of the face using a sequence of convolution layers, pooling layers and inception modules. During user sessions, this algorithm analyzes dynamic videos collated by the users webcam during user involvement and engages such movements as eye movement, brow shape, mouth positioning. The model also detects such features as a slight smile or a furrowed brow, which are some of the subtle features of an individual's emotional state; these relate to specific emotions, such as happiness, sadness or anger. The result is that the

platform is able to understand precisely what the emotional state of the user is and modify its feedback accordingly. The ability of Inception v3 to perform these tasks with correct processing opens new horizons, it is definitely suitable for a fast-paced, real-time in-getting to know the environment

Facts analysis with Voice:

The evaluation of information with the use of voice is concerned with the complicated and advanced computation techniques employed in processing and understanding customers' voice recordings. When collected, voice recordings are converted first to a spectrogram or a feature vector that retains essential noise characteristics such as pitch, tone, and intensity. These features are examined in order to identify trends that are connected to certain emotional states. A major pitch is likely to be related to pleasure or joy, while a low or flat intonation of the voice may be interpreted as sadness or boredom. For instance, machine learning algorithms can be trained to perform this task on audio features to detect the subtle changes in emotions in a learner during interaction. Thus providing the system with more information on the emotional dynamics of the learner and reasons for providing such feedback and support.

Information analysis with Face:

The ability to analyze facial emotions is crucial in providing an in-depth understanding of the user's emotional state. With this module, facial recognition technology is employed to capture and analyze expressions—such as the action of the eyes, mouth, and brows—in real-time. These expressions have been analyzed to identify emotions such as happiness, sadness, anger, astonishment, and fear. The analysis of the head movement will also be combined with the voice analysis, so that a total emotional map of the client would be utilized. This allows the platform to modify its internal responses speaking- for example, congratulating, changing the levels of difficulty or offering techniques to control pressure- with respect to both speech and facial expressions.

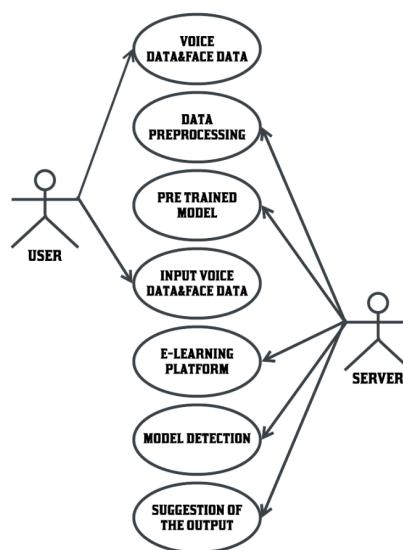


Fig 1.2 Use Case diagram

VI. SYSTEM ARCHITECTURE

The architecture includes Data Preprocessing (noise reduction and normalization), Feature Extraction using CNN for spatial features, and LSTM for temporal data analysis. Detected emotional states from audio or facial features guide the system in delivering Adaptive Feedback in real time, promoting emotional well-being and enhancing engagement. This layered structure enables efficient, responsive, and emotionally adaptive interaction for users in a learning setting providing tailor-made suggestions or motivational prompts.

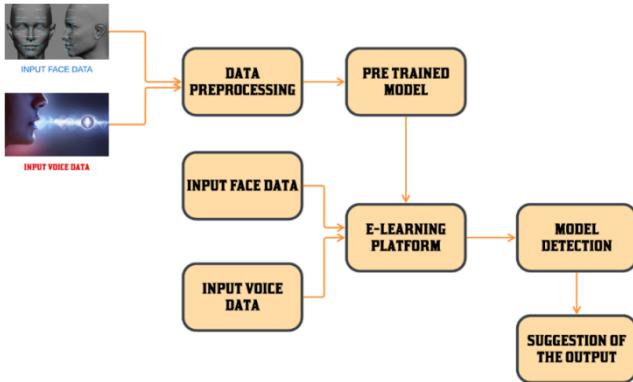


Fig 1.3 Architecture diagram

VII. RESULTS & DISCUSSIONS

Measurable gains in user engagement, versatility, and emotional support within the learning environment have emerged from the use of the emotion-driven e-learning platform. The platform's capacity to identify and react to human emotions, including joy, sorrow, rage, and frustration, has been assessed through iterative testing and suggestions, producing a number of noteworthy results.

Accuracy of Emotion identification: During controlled experiments using labeled audio data, the CNN-LSTM model showed excellent performance in emotion identification, recognising primary emotions like happy, sorrow, and rage with an accuracy of about 90%. However, because of the subtleties in tone, it demonstrated somewhat less accuracy in identifying more subdued emotions, including moderate annoyance or neutral moods. The efficacy of the emotion detection method is still confirmed by the model's overall performance, which combines CNNs for feature extraction (such as tone and pitch) and LSTMs for identifying temporal patterns. Because of its precision, the platform can offer contextually appropriate feedback, making the user experience more responsive and encouraging.

User Satisfaction and Engagement: According to user input, there was a notable increase in both of these metrics, especially when the system modified its replies in reaction to emotional cues. Users who exhibited grief or dissatisfaction, for instance, were given motivational cues, which increased their focus and strengthened their will to learn. The platform's capacity to create

a supportive environment was demonstrated by the fact that users were more likely to stick with their sessions when they received sympathetic feedback during trying times. According to these results, emotional intelligence in e-learning systems can significantly improve user happiness by encouraging motivation and emotional health, which in turn promotes more persistent learning behaviors.

Impact on Learning results: Learning results were enhanced by the platform's adaptive feedback, which also enhanced emotional involvement. The technology modified the pace of content distribution or condensed explanations for users displaying signs of frustration so they could take in information more comfortably. Improved memory and comprehension resulted from this individualized approach, especially in complicated subjects where emotional difficulties like stress or confusion are frequent. The capacity to adapt information delivery in real time in response to emotional input shows how important emotional intelligence is for improving understanding and learning efficacy, which in turn promotes more effective and efficient learning.

Limitations and Difficulties: Despite the encouraging results, a number of difficulties were identified. The inability of the system to discern small emotional distinctions was highlighted by the fact that it occasionally misclassified emotions with similar aural cues, such as moderate annoyance from neutral tones. Furthermore, the effectiveness of emotion identification was occasionally impacted by background noise in user contexts, especially in less controlled conditions. Emotions outside of the main categories, including excitement or boredom, which could improve the system's responsiveness, were similarly difficult for the model to handle. In order to accommodate a wider range of emotional states, these difficulties imply that the system would profit from advancements in noise-filtering methods as well as an extension of its emotional detection range.

Technical Performance: The CNN-LSTM model processed audio inputs effectively and provided feedback almost instantly, usually within two seconds per input. The platform's technical performance was generally strong. Sustaining a smooth and engaging user experience requires this low latency. Nevertheless, sporadic latency during concurrent interactions was noted, indicating potential for improvement to better manage concurrent user sessions. Consistent response times should be the main goal of future performance enhancements, especially as the platform grows to support more users.

User Input on System Usability: According to surveys and interviews, users were quite pleased with the platform's emotional reactivity and ease of use. A sense of connection and support was facilitated by the fact that many participants reported feeling "understood" by the system. This is in line with the platform's objective of establishing a comprehensive learning environment that attends to both emotional and cognitive demands. Users recommended enhancements including broadening the emotional states the system can identify and improving feedback personalisation to better accommodate particular emotional expressions. These recommendations reaffirm the necessity of ongoing improvement and customisation to raise the platform's emotional intelligence.

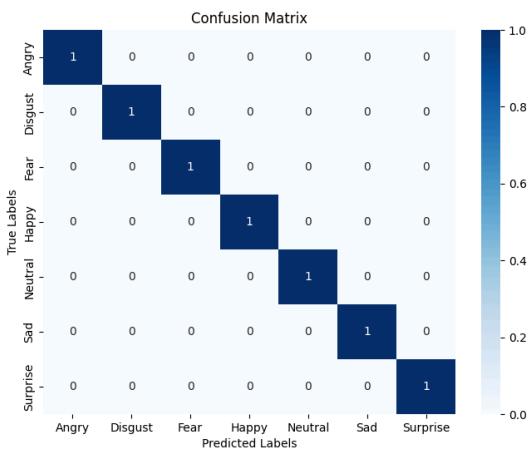


Fig 1.4 Confusion matrix

Discussion of Future Improvements: Future research could concentrate on making the emotion detection model more sensitive to more complex emotional states, such as excitement or boredom, in order to solve the issues found and enhance system performance. By broadening the emotional spectrum, the system's general adaptability to a variety of learning situations would be enhanced. Furthermore, incorporating noise-canceling methods may enhance the precision of emotion identification in many contexts, guaranteeing a more uniform user experience. Using predictive analytics, where the system may use past user data to forecast future emotional states and modify learning courses accordingly, is another exciting avenue. This proactive change may provide even more emotional support and personalisation, increasing the platform's capacity to promote a positive and effective learning environment.

To sum up, the emotion-driven e-learning platform has a lot of potential to enhance learning outcomes, emotional support, and user engagement. Notwithstanding certain difficulties, incorporating emotional intelligence via facial and speech recognition is a big step in the direction of developing a more adaptable and sympathetic educational environment. The platform has the potential to completely transform online education by providing individualized, emotionally sensitive learning environments that foster both cognitive and emotional development, provided that its emotional detection skills are further improved and expanded.

VIII. FUTURE ENHANCEMENTS

Future improvements to the Emotion-Driven Voice and Face Interaction System for User Motivation and Support have enormous potential to improve the user experience and further optimize its performance. The incorporation of a wider variety of emotional states is one important improvement. The platform would be able to identify and react to a greater range of emotional states if it included more subtle emotional cues, such as mild worry, contentment, or perplexity. At the moment, the system can identify core emotions like happiness, sadness, and irritation. With this extension, the system would be able to

give more contextually relevant and nuanced feedback, which would enhance its capacity to more accurately and sympathetically respond to users' emotional needs.

Adding adaptive machine learning models, which continuously learn from user interactions, is another possible direction for development. The platform may utilize reinforcement learning techniques to improve its answers and strategies in response to user feedback as users interact with the system over time. As more information about the user's emotional preferences and reactions is gathered, the system's emotional responses will also change, becoming more accurate and personalized in a cycle of continual improvement. This kind of flexibility would guarantee that the platform stays appropriate for the user's emotional state, enhancing effectiveness and engagement all the way through the learning process. Moreover, adding multi-modal inputs to the system might offer a more thorough comprehension of the user's emotional condition. An additional layer of emotional insight could be added by using data from wearable devices (such as skin conductance sensors or heart rate monitors) in addition to voice and facial expression analysis. The technology would be able to identify emotional states like stress, relaxation, or slight mood swings that might not be fully conveyed by speech or facial expressions alone by comparing verbal signals with physiological data. This multifaceted strategy would improve the system's real-time emotional state assessment capabilities, resulting in more precise, fast, and flexible feedback.

These improvements would boost the platform's capacity to provide users with individualized, sympathetic, and comprehensive help in addition to improving its emotional reactivity. The Emotion-Driven Voice and Face Interaction System has the potential to greatly improve user engagement and overall learning results by developing further in tandem with advances in machine learning and multi-modal data integration.

IX. CONCLUSION

To conclude, the incorporation of superior emotional popularity through voice and facial features evaluation into the e-studying platform marks a widespread step closer to enhancing the learning revel in by making it extra personalized and tasty. By understanding the emotional states of learners, the platform can provide actual-time, custom designed feedback and help, addressing emotional desires as they stand up. This emotional sensitivity is not most effective and creates a more interactive and motivating surroundings however it also aids novices in overcoming boundaries, boosting their confidence and improving ordinary educational results. In essence, this initiative highlights the fee of mixing era with emotional intelligence, developing an extra responsive, compassionate, and impactful mastering revel in. Through this innovation, the platform has the capability to reshape the manner college students have interaction with learning, fostering more potent connections to their research even as promoting each emotional nicely-being and academic fulfillment.

X. REFERENCES

- [1] N. Dhariwal, S. C. Akunuri, Shivama and K. S. Banu, "Audio and Text Sentiment Analysis of Radio Broadcasts," in IEEE Access, vol. 11, pp. 126900-126916, 2023, doi: 10.1109/ACCESS.2023.3331226.
- [2] The Analysis of Music Emotion and Visualization Fusing Long Short-Term Memory Networks Under the Internet of Things
- [3] K. Zaman, M. Sah, C. Direkoglu and M. Unoki, "A Survey of Audio Classification Using Deep Learning," in IEEE Access, vol. 11, pp. 106620-106649, 2023, doi: 10.1109/ACCESS.2023.3318015
- [4] A. Dutt and P. Gader, "Wavelet Multiresolution Analysis Based Speech Emotion Recognition System Using 1D CNN LSTM Networks," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 2043-2054, 2023, doi: 10.1109/TASLP.2023.3277291.
- [5] H. -D. Le, G. -S. Lee, S. -H. Kim, S. Kim and H. -J. Yang, "Multi-Label Multimodal Emotion Recognition With Transformer-Based Fusion and Emotion-Level Representation Learning," in IEEE Access, vol. 11, pp. 14742-14751, 2023, doi: 10.1109/ACCESS.2023.3244390.
- [6] M. Xu, F. Zhang, X. Cui, and W. Zhang, "Speech emotion recognition with multiscale area attention and data augmentation," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Jun. 2021, pp. 6319–6323.
- [7] C. Gong, D. Wang, M. Li, V. Chandra, and Q. Liu, "KeepAugment: A simple information-preserving data augmentation approach," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 1055–1064.
- [8] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, "RandAugment: Practical automated data augmentation with a reduced search space," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2020, pp. 3008–3017.
- [9] A. Dang, T. H. Vu, L. D. Nguyen, and J.-C. Wang, "EMIX: A data augmentation method for speech emotion recognition," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Jun. 2023, pp. 15
- [10] B. T. Atmaja and A. Sasou, "Effects of data augmentations on speech emotion recognition," Sensors, vol. 22, no. 16, p. 5941, Aug. 2022.
non-intrusive assessment methods(literature survey).
- [11] Kumar P; Salman Latheef T A; Santhosh R(2023), "Face Recognition Attendance System Using Local Binary Pattern Algorithm," 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN), Vellore, India, 2023,pp.1-6, doi: 10.1109/ViTECoN58111.2023.10157843.
- [12] P. Kumar, S. Senthil Pandi, T. Kumaragurubaran and V. Rahul Chiranjeevi (2024), "Human Activity Recognitions in Handheld Devices Using Random Forest Algorithm," 2024 International Conference on Automation and Computation (AUTOCOM), Dehradun, India, 2024, pp. 159-163, doi: 10.1109/AUTOCOM60220.2024.10486087.

RE-2022-418506-plag-report

ORIGINALITY REPORT



PRIMARY SOURCES

1	www.ijisae.org Internet Source	1 %
2	elibrary.stipram.ac.id Internet Source	1 %
3	www.scilit.net Internet Source	<1 %
4	Submitted to University of North Texas Student Paper	<1 %
5	link.springer.com Internet Source	<1 %
6	Submitted to Advance HE Student Paper	<1 %
7	Submitted to Cardiff University Student Paper	<1 %
8	docshare.tips Internet Source	<1 %
9	Khalid Zaman, Melike Sah, Cem Direkoglu, Masashi Unoki. "A Survey of Audio	<1 %

Classification using Deep Learning", IEEE

Access, 2023

Publication

-
- 10 Naman Dhariwal, Sri Chander Akunuri, Shivama, K Sharmila Banu. "Audio and Text Sentiment Analysis of Radio Broadcasts", IEEE Access, 2023 <1 %
Publication
-
- 11 acceleron.org.in <1 %
Internet Source
-
- 12 etd.auburn.edu <1 %
Internet Source
-
- 13 www.mdpi.com <1 %
Internet Source
-
- 14 Samira Hazmoune, Fateh Bougamouza. "Using transformers for multimodal emotion recognition: Taxonomies and state of the art review", Engineering Applications of Artificial Intelligence, 2024 <1 %
Publication
-
- 15 Yujing Cao, Jinwan Park. "The Analysis of Music Emotion and Visualization Fusing Long Short-Term Memory Networks Under the Internet of Things", IEEE Access, 2023 <1 %
Publication
-

Exclude quotes

On

Exclude matches

Off

Exclude bibliography

On

CO-PO Mapping

PROJECT WORK COURSE OUTCOME (COs):

CO1: On completion it will prove as a major breakthrough in digital transformation of college management leveraging end-to-end technologies.

CO2: It will ease out the management overhaul and boost better transparency and robustness to the entire setup.

CO3: Given the huge amount of data available in the educational sector, especially in the colleges, technologies like Machine Learning and AI can be used to increment student performance and job-market ready.

CO4: It helps in keeping the entire system snappy and ensures all endpoints are taken care of, reducing the overall waiting periods in the traditional working.

CO5: Students will be able to publish or release the project to society.

PROGRAM OUTCOMES (POs)

PO1: Engineering Knowledge: Apply the knowledge of engineering fundamentals, mathematics, science and technology and an engineering specialization to the solution of complex engineering problems.

PO2: Problem analysis: Ability to apply deep learning methodologies to solve computational tasks, model real world problems using appropriate datasets and suitable deep learning models. To understand standard practices and strategies in software project development using open-ended programming environments to deliver a quality product.

PO3: Design/development of solutions: Design solution for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety.

PO4: Conduct investigations of complex problems: Use research - based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis the information to provide valid conclusions.

PO5: Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6: The Engineer and society: Apply reasoning informed by the contextual knowledge to assess social, health and safety issues and the consequent responsibilities relevant to the professional engineering practice.

PO7: Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental context, and demonstrate the knowledge of, and need for sustainable development.

PO8: Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practices.

PO9: Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10: Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11: Project management and finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO12: Life-long learning: Recognize the need for and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAM SPECIFIC OUTCOMES (PSOs):

PSO1: Foundation Skills: Ability to understand, analyze and develop computer programs in the areas related to algorithms, system software, web design, deep learning and cloud computing for efficient design of computer-based systems of varying complexity. Familiarity and practical competence with a broad range of programming languages and open-source platforms.

PSO2: Problem-solving Skills: Ability to apply mathematical methodologies to solve computational tasks, model real world problems using appropriate data structure and suitable algorithms. To understand standard practices and strategies in software project development using open-ended programming environments to deliver a quality product.

PSO3: Successful Progression: Ability to apply knowledge in various domains to identify research gaps and to provide solutions to new ideas, inculcate passion towards higher studies, creating innovative career paths to be an entrepreneur and evolving as an ethically responsible computer science professional.