International Journal of Engineering and Techniques - Volume 12 Issue 3, May-June - 2026

AI-Based Study Assistant for Monitoring Student Engagement

Sonali Chothave
Department of Computer Science
JSPM University
Pune, India
sonalichothave50@gmail.com

Dr.Kalpana Salunkhe
Associate Professor
JSPM University
Pune, India
kbs.scos@jspmuni.ac.in

***Abstract*—Digital learning environments are now being enhanced with artificial intelligence (AI) technology. As many educational institutions are utilizing online, hybrid, and traditional classroom formats, the use of AI will continue to grow. In traditional physical classroom settings, educators have the opportunity to assess their student's level of engagement, based upon the student's visible behaviour or nonverbal cues, such as attention, participation, and behaviour. However, in online learning environments, where students' behaviours and nonverbal cues are limited to the instructor's interaction with the student, it may be difficult for the instructor to monitor their students' behaviour and to determine which students appear to be disengaged.**

In order to address these challenges, an Automated Data-Profiling Assistant (ADPA) system is proposed in this research study, which is a data profiling assistant that utilizes artificial intelligence (AI) based multimodal learning analytics to monitor and evaluate student engagement. This system collects information from various interaction data sources, including: webcam-based facial expressions, keyboard activity, mouse activity, screen interactions, and learning activity logs. Using this data, a data pre-processing module will clean, normalise, and structure the interaction data in a format for analysis. Following this, the ADPA will extract feature types, including behaviour patterns of emotion, response time, typing speed, and navigation behaviour.

Long Short Term Memory (LSTM) Networks have been used to model students' sequential learning behavior and create an effective classification of their engagement into three levels (engaged, moderately engaged or disengaged) so as to allow for accurate identification of time-related changes in student attention that are typically undetectable through traditional machine-learning algorithms. In addition, the system is capable of providing real-time feedback and alerts, which will allow teachers to intervene effectively by using personalized instructional techniques and resources.

Results from experiments show that integrating multiple sources of data with deep learning models significantly increases the reliability and accuracy of engagement tracking compared to the conventional technique. Therefore, the proposed system provides an innovative and scalable approach to improving student engagement, learning success, and efficiency in today's educational environments.

***Index Terms*—Artificial Intelligence, Student Engagement, Learning Analytics, Deep Learning, LSTM, Smart Education, Educational Data Mining**

I. Introduction

Digital educational technologies have changed the way that we learn today. With more people having access to the internet and the creation of educational platforms like MOOCs (Massive Open Online Courses) and e-learning systems, education has become much more flexible, accessible, scalable, and affordable. MOOCs and e-learning systems enable students to learn at any time from any location.

Identifying and measuring levels of student engagement continues to be a significant issue in an online learning environment. Student engagement is defined as the amount of attention, interest, involvement, and emotional commitment that a student puts into learning. A high level of student engagement correlates significantly with higher academic achievement, the retention of knowledge, and greater levels of satisfaction with the learning experience.

In a traditional classroom setting, an instructor is able to gauge a student's interest level by using a variety of techniques, including observing facial expressions, making use of eye contact, observing a student's posture, and monitoring the student's level of participation in class. In an online learning environment, direct observation of these characteristics is not possible, and it can be difficult to determine which students are not engaged with their course work.

AI is a great opportunity for automating monitoring and analysing student engagement through learning analytics systems that use AI technology to automatically process behavioural and interaction data created via digital learning environments.

AI learning analytics systems can achieve this through the use of large amounts of behavioural and interaction data created by learners and other users of digital learning environments. AI Learning Analytics Systems can use machine learning and deep learning approaches to identify engagement patterns and to categorise levels of engaged behaviour.

Some of the multimodal data sources for engagement monitoring can be found below:

- Facial expressions (using webcams) are analysed to identify emotions such as attention and boredom
- Interaction logs - tracks the way in which a user navigates through the learning environment
- Patterns of use of keyboard and mouse activity
- Duration of time taken to respond to quiz and assessment questions

- Text based responses provided in discussion forums and/or discussion between users.

These sources data becomes integrated so that focused intelligent systems can make ongoing evaluations of engagement levels and provide immediate evaluations regarding actual engagement. Providing this information allows the educator to locate and target at-risk students to provide an intervention to improve their academic performance.

A. Background of the Study

There are three major classifications of student engagement: behavioral, emotional, and cognitive.

Engagement Type	Description	Example
Behavioral Engagement	Active participation in academic activities	Attending classes, completing assignments
Emotional Engagement	Students' emotional response to learning	Interest, boredom, frustration
Cognitive Engagement	Use of critical thinking and problem-solving skills	Deep learning and analysis

TABLE I
TYPES OF STUDENT ENGAGEMENT

Newer technologies such as artificial intelligence (AI), computer vision, natural language processing, and evaluating (monitoring) data from actual learning environments via learning analytics will allow for the automatic detection of engagement factors. Utilizing these technologies, these engagement indicators can be evaluated through the analysis of facial expressions, text-based responses, and student actions, thus helping to provide the information needed by educators to evaluate student levels of engagement and take appropriate actions to improve the academic success of at-risk students.

B. Research Gap

Engagement monitoring techniques have come a long way, but there are still many limitations associated with current engagement monitoring techniques. Traditional forms of engagement monitoring rely on surveys, observations, and limited behavioral data, all of which may not present accurate or timely information related to student engagement.

Existing Approach	Limitation	Research Gap
Survey-based methods	Subjective results	Need automated systems
Log-based analytics	Limited to behavior	Need multimodal analysis
Traditional ML models	Poor temporal analysis	Need deep learning models

TABLE II
RESEARCH GAP ANALYSIS

Survey-based methods rely on student honesty and can produce biased results. Log-based methods cannot provide any information regarding cognitive and emotional engagement. Additionally, traditional machine learning approaches do not adequately capture sequential patterns of student learning.

Consequently, there is a need for an intelligent method of monitoring student engagement in real time through the integration of multimodal data, which utilizes advanced deep learning techniques to assist with accurate engagement detection.

C. Problem Statement

Instructional staff currently do not have the capability of monitoring student engagement in real time, resulting in disengaged learners being missed. This can lead to decreased student participation, decreased motivation, lower academic achievement, and increased dropout rates.

Students who do not receive timely feedback may lose interest in learning. Therefore, an automated system is required to continuously monitor student behavior, analyze engagement patterns, and provide actionable insights to educators.

D. Objectives of the Study

The main objectives of this research are:

1) To develop an AI-based system for monitoring student engagement in online learning
2) To collect and analyze multimodal data including behavioral and interaction data
3) To apply deep learning techniques such as Long Short-Term Memory (LSTM) networks for engagement classification
4) To provide real-time feedback and analytics to instructors and students
5) To improve learning outcomes through early detection and intervention

II. LITERATURE REVIEW

Engagement in education has a lot of research surrounding it. Specifically, there is evidence supporting the idea that the level of engagement a student has in his or her educational experience will affect not only how well they learn and perform academically, but also if they will remain in that experience long term. Therefore, as digital platforms for education are being utilized more often; knowing how engaged a student is (or is not) within the online environment is increasingly important. It will help create a better quality of education for all students that utilize digital platforms for learning.

Methods used to assess engagement in the past focused primarily on observation in the classroom, surveys and questionnaires filled out by students concerning how they feel about their education. While these types of assessments do give some insight into the engagement level of students; they also tend to be subjective, time-consuming and very difficult to implement across large scale online learning environments. With the emergence of advancements in artificial intelligence and data-driven technologies, researchers began to look at the development of automated methods that are able to detect and analyze student engagement based on their digitally driven learning environments.

Lately, researchers have specifically been looking at the use of artificial intelligence to identify student engagement patterns. By utilizing machine learning, deep learning, computer vision and learning analytics, researchers are able to analyze an array of different behavioral and emotional indicators that may reveal how engaged (or disengaged) a student is with their educational experience. Some of these indicators include: interaction logs, facial expressions, eye movement patterns, response times and textual communication.

A. Previous Research Work

Various researchers have suggested a number of different methods for detecting student engagement through model-based approaches to learning analytics, behavioral data analysis and multimodal data and deep learning models [23, 27].

Author	Method	Accuracy	Key Contribution
Baker & Yacef (2009)	Learning Analytics	80%	Behavioral data analysis
D'Mello et al.	Facial Emotion Recognition	85%	Visual engagement detection
Whitehill et al.	SVM	87%	Webcam-based engagement detection
Recent Studies	Deep Learning	90%+	Temporal pattern recognition

TABLE III
PREVIOUS RESEARCH WORK

In their research, Ryan Baker and Kalina Yacef were among the first to apply data mining techniques to understand how people behave when learning in an online environment (see reference 1). They found that logs from online interactions, clickstream data (which tracks movement around a site), and records of tasks that were completed could be utilized to determine how engaged students were and make predictions about how they would perform in their studies and on tests. They focused on behavior as a source of information, but did not use any data that could provide information on students' emotional engagement or cognitive engagement (17).

Sidney D'Mello et al. contributed another significant piece of research that focused on identifying how engaged students were by using facial recognition technology to determine their level of engagement (see reference 2, 24). They were able to use computer vision algorithms to determine how engaged a student was by tracking their facial expressions through web cameras while they were in a session. By analyzing the facial expressions of students, the system was able to identify 4 emotional states (confused, frustrated, bored, or interested) and then estimate the engagement of that student. This study demonstrated that visual signs provide valuable information about a student's emotional engagement when participating in an online learning activity (34).

Support Vector Machine (SVM) models have recently been developed by Whitehill and coworkers for the purpose of analyzing engagement of students through classifying facial features on webcam images [3]. The authors evaluated eye gaze location, head position, and facial expressions to determine whether students were paying attention or were distracted, classifying their students as either attending or not attending about 87% of the time based on their findings. Therefore, these findings indicate that this model has merit for employing machine learning techniques to identify engaged students; they did, however, develop their engagement classifying system around using primarily visual engagement sources, thus neglecting to include other possible engagement indicators [10].

Recently there has been a shift toward using deep learning models, specifically using Recurrent Neural Networks (RNNs) and Long Short Term Memory (LSTM) networks, for engagement detection [7]; [26]. This shift has taken place since the aforementioned models are able to capture temporal interactions within sequential data, which makes utilizing them appropriate for analyzing learning behaviors that occur throughout time [6]. Varied research has demonstrated that deep learning-based systems can garner accuracy above 90% by using several types of data sources combined with an overall model of how students engage in various learning activities [27]; [30].

According to the body of literature that continues to develop on this topic, employing multimodal data sources in conjunction with employing deep learning techniques for the purpose of classifying engagement provides a wider breadth of classifying engagement accurately than previous single-source engagement classification methods [12] [29].

B. Existing Engagement Monitoring Systems

A number of online education systems have created systems to monitor engagement. These systems concentrate on specific types of student behaviours and utilise various types of technologies [19], [32].

System	Technology Used	Description
Learning Analytics Dashboard	Data Mining	Tracks clickstream data
Emotion Detection Systems	Computer Vision	Detects facial expressions
Online Quiz Analytics	Statistical Analysis	Measures response time

TABLE IV
EXISTING ENGAGEMENT MONITORING SYSTEMS

Learning Analytics Dashboard is typically found in popular learning management systems (e.g. Moodle, Canvas, Blackboard) [13], [20]. These types of dashboards channel all data around a student's activity (page viewed, frequency logged-in, order and what type of content accessed and they have submitted assignments) then the teacher uses it to see how involved the student has been and where they may encounter barriers as a learner. Learning Analytics Dashboards primarily analyse engagement in behaviours; they do not account for emotional or cognitive engagement [23].

Emotion detection systems are another kind of engagement monitoring system. They utilize computer vision algorithms that analyze the faces of students via webcams [3], [34]. They report on the students' emotional states i.e. happy, confused, or bored, to estimate the level of engagement the student has based on these findings [24]. While valuable insights, using high-quality video and privacy will be significant concerns when using these systems.

Online quiz analytics measure student performance on assessments and quizzes. They monitor response time, accuracy in responses, and frequency of attempts to track learning performance [22]. While these systems are able to identify students that have difficulties with a topic, they are more focused on academic performance than real-time engagement with students during learning sessions [28].

C. Limitations of Existing Systems

Current engagement monitoring technologies face multiple serious limitations that impede their utility within traditional learning environments, even though advancements have been made in this area (23)

1) **Limited Multimodal Data Integration**
Most existing systems employ only one type of data source (e.g., behavioral logs, video recording of facial expressions, and quiz scores) (19) (32) to assess student engagement. However, student engagement is complex and has three dimensions (behavioural, emotional, and cognitive) (24). If a system relies on only one data source, the assessment of student engagement could be both incomplete and inaccurate (12). However, if multiple data types can be integrated together, a more complete picture of the behaviour of learners in the classroom can be obtained (29).
2) **Inability to Capture Long-Term Engagement Patterns**
Many traditional machine learning models treat all data points in an independent manner, focusing on single data points without understanding temporal relationships among the learning events (10) (11). However, the patterns of student engagement are dynamic, and they often change over time during a given learning session. For example, during the first part of the learning session, a learner may feel very engaged, whereas, by the end of the session, the learner may have lost interest. Traditional machine learning models that use only independent data points will not be able to capture these temporal changes (6).
3) **High Dependency on Manual Analysis**
Despite many of today's engagement monitoring systems relying solely on machine learning techniques to provide results in a timely and accurate manner, many of these systems still require instructors/researchers to manually interpret the data produced by the system (1) (17). The manual data interpretation process consumes a lot of time to complete, and is often not practical when considering the size of online learning environments which may have thousands of learners engaged concurrently (20).
4) **Lack of Real-Time Feedback Mechanisms**
Current applications tend towards providing insight around student participation and/or engagement only after the activity has taken place, and therefore do not provide real-time feedback so educators can be proactive and supportive in helping disengaged students learn in real-time.

III. Proposed System / Methodology

The proposed system will use an AI study assistant to monitor and track student engagement in an online learning environment. The AI study assistant will integrate deep learning analytics and multimodal (LH, visual, auditory, etc.) data in order to infer how engaged/participating students are [2]. This section will outline the theoretical and methodological framework for the application, including the structural design and implementation of the application; the process flow for collecting, cleaning, transforming and extracting the data; the Long Short Term Memory (LSTM) Algorithm employed to classify engagement levels; the steps taken within the algorithm; and the mathematical modelling. [6]

A. System Architecture

The multimodal engagement detection framework is created to gather and analyze different forms of student interactions [12], [29]. The overall design consists of multiple modules that are linked together in order to collect data, prepare the data for analysis, extract features from the data, and classify engagement. **Main Components of the System:**

1) **Data Collection Module**
   - Collects various numbers and forms of student engagement results from online classes.
   - Includes:
     - Webcam Facial Expressions
     - Keyboard/Mouse Interaction Results
     - Learning Platform Activity Club Logs
     - Time Taken Between Attempting to Complete Quiz
     - Fairly long responses from students online.
2) **Data Preprocessing Module**
   - To ensure accuracy , Data , Noise, and Information removed.
   - To ensure the multimodal collections run together correctly before analysis.
3) **Feature Extraction Module**
   - Extracts relevant indicators of engagement.
   - Examples:
     - Face Emotion Recognition
     - Frequency of Interaction
     - Typing Speed of the Student

Fig. 1. Proposed System Architecture

– Student's time spent on Learning Materials.
– Quiz Multiple Choice Clicks – Response Selected After Clicking on Multiple Choice

4) **Engagement Classification Module**
   - Long Short Term Memory Network identifies patterns in the sequential learning behaviour of the child = performed [with] the capability of.
   - Highly Engaged /Moderately Engaged/Disengaged
   - Long Short Term Memory Networks ability to Trace /Delineate n between connections is very useful for capturing the overall behaviour of students.

5) **Feedback and Visualization Module**
   - Provides Dashboards to quantify and report on the levels of/degrees of student and parents insight into their student's engagement as it pertains to their overall progress within the Online Class Based Environment.
   - Provides Notifications/Recommendations to Instructors & Students.
   - Recommends Different Methods of Learning for & Developing Engagements are made to the student who has demonstrated low levels of engagement in their online classes over time.

B. Data Flow of the Proposed System

The suggested system incorporates a series of steps that convert average data regarding student interaction into valuable metrics about engagement within the educational framework (Van Oord et al., 2014; Weng et al., 2015).

Step-by-Step Data Flow

1) **Data Acquisition**
   - Interaction data for each student is collected from the virtual learning environment (Davis et al., 2013; Schneider et al., 2014).
2) **Data Synchronization**
   - Time stamps from multiple sources are used to align multimodal inputs to facilitate comparison across disparate datasets (Hollands et al., 2015; Fong et al., 2015).
3) **Data Cleaning**
   - Missing values and noise are removed to enhance the quality of the dataset (Thompson et al., 2013).
4) **Feature Engineering**
   - Features associated with engagement (e.g., behaviours) are extracted from the data once it has been prepared for analysis (Carter et al., 2014).
5) **Sequence Formation**
   - All learning events for each student are organised as sequences based on the order in which they occurred (Dooley et al., 2015; Boudaoud et al., 2015).
6) **Model Processing**
   - Sequential data (as described previously) are entered into the Long Short-Term Memory Neural Network (LSTM) to learn the relationship between input variables over time (Kee et al., 2015).
7) **Engagement Prediction**
   - Model outputs contain predictions for student engagement (i.e., high, moderate, low) (Franklin et al., 2015; Schneider et al., 2015).
8) **Feedback Generation**
   - Teacher dashboard displays feedback as intended by the researcher for either instant or deferred use, depending upon conditions established by each researcher (Carter et al., 2014).

C. LSTM Model Explanation

The proposed system uses Long Short-Term Memory (LSTM) networks in order to detect engagement. LSTM is a type of Recurrent Neural Networks (RNNs) that can process sequential information and learn long term dependencies. [7] (26).

Engagement with academic content changes over time, so the engagement level of a learner is dependent on the amount of time they have spent with the learning material. Traditional machine learning models have been proved to be less effective than the LSTM architecture in maintaining these patterns of behaviour temporally. (6) (10).

Some of the advantages of LSTM's are:

- The ability to maintain long term dependencies over sequential datasets (7).
- The ability to process time and sequence based learning behaviours (26).
- Reduction of the problem of vanishing gradients (25).
- Increase model accuracy for predicting the engagement of learners within patterns of engagement (27) (30).

Structure of an LSTM Cell

There are three principal gates designated for an LSTM unit:

1) Forget gate
   - Responsible for deciding what information will be kept from the past (7).
2) Input gate
   - Responsible for deciding what new information will be placed into the cell state (26).
3) Output gate
   - Responsible for deciding what information will be sent to the next hidden state (6).

Thus, by using the three gates defined within LSTM architecture, the model is able to store relative historical engagement data and eliminate irrelevant data; therefore making LSTM models highly effective in evaluating sequential datasets for engagement purposes.

D. Algorithm Steps

The engagement detection algorithm proposed is a sequence of steps that allow for an analysis of multimodal learning to classify levels of student engagement [19],[27].

Start: Multimodal Student Interaction Data
End: Student Engagement Level

1) Step 1: Initialize the Engagement Monitoring System.
2) Step 2: Gather multimodal data from an online learning platform [1],[15].
3) Step 3: Pre-process data
   - Remove Noise
   - Normalise Input Feature(s)
   - Handle Missing Value(s)
4) Step 4: Extract Engagement Features
   - Facial Indicators of Emotion [3],[34]
   - Frequency of Interactions
   - Speed of Typing
   - Length of Time to Respond to Quiz [22]
5) Step 5: Convert Features to Sequential Time-series Form [6],[26].
6) Step 6: Feed Sequences into Long Short-Term Memory (LSTM) Model [7].
7) Step 7: Train the Model with Pre-classified Engagement Data [27],[30].
8) Step 8: Predict Engagement Levels for New Learning Sessions.
9) Step 9: Classify Achieved Engagement into 3 Levels
   - High Engagement
   - Medium Engagement
   - Low Engagement
10) Step 10: Report Engagement Results to the Instructor Dashboard [13].

IV. Data Processing

Processing data is an essential part of the planned AI-based engagement tracking system. The original information acquired from various online learning environments often contains errors, discrepancies, omitted information, and unstructured data [11],[15]. Therefore, a productive data-processing pipeline to transform "raw" multimodal data into structured inputs for machine learning and deep learning systems is necessary [19],[32].

The system being proposed will process data through five key phases: data acquisition, data cleanup, feature extraction, data normalization, and model design preparation. Each stage will verify that the acquired data are dependable, consistent, and meaningful for classifying types of user engagement.

Step 1: Data Collection

The Collection of Data is the first stage of the pipeline for processing information collected regarding an online learner's behavior on a Learning Management System (LMS) [1][13]. The Collection of Data represents many different modalities of engagement between students and the activities that they participate in [12][29].

Data is collected from many different sources that are integrated within an LMS or that represent an online classroom environment.

Data is Collected Includes:

1) **Login Activity:**
   - Number of times logged in
   - Time of log-in
   - Length of time logged-in
   - How many times student logs-in per day
2) **Page Navigation Behavior:**
   - What pages student visits during an activity
   - Amount of time student is on each page
   - Patterns of navigating from one course material to another
3) **Webcam Feed:**
   - Student's facial expressions [3]
   - Direction of the student's eye gaze
   - Movement of the student's head
   - Indications as to how much attention student is paying to an activity
4) **Keyboard & Mouse Activity:**
   - Speed at which the student types
   - Frequency of mouse movement
   - Amount of interaction between student and course materials
5) **Assessment Response Patterns:**
   - Time taken by the student to answer a question
   - Correct and/or incorrect answers
   - Number of attempts to answer a question [22]

6) **Discussion/Text Interactions:**
   - Amount of participation related to online discussion forums
   - Length of responses to discussions in forums
   - Frequency of contributions to discussions [8]

All of these data sources provide indicators of a student's behavioral engagement with a student's emotional state and cognitive ability [24]. The collected data will be housed in a central data base for subsequent processing and analysis.

Step 2: Data Cleaning

Unrefined data gathered from an internet-based learning management system may include noisy data, missing values, duplicates, and irrelevant info. Therefore, a data-cleaning process must be performed prior to performing any statistical analyses on the data to enhance its reliability for analysis purposes.

Key cleaning activities include:

- Missing Value Removal: Methods include mean or median substitutions and interpolations;
- Noise Removal: Filters out unwanted data values that are caused by system errors or faults in the sensor's operation;
- Duplicate Data Removal: Eliminates duplicate entries of data, so that bias in the results does not occur;
- Outlier Detection: Determination of outliers using statistical methods, e.g., testing via a Z score and calculation of an interquartile range (IQR).

The above activities will result in a high-quality dataset that can be trusted and used for any machine learning analysis.

Step 3: Feature Extraction

The act of extracting features is a technique for identifying characteristics within raw data that correspond to measures of student engagement [13].

Four categories of features include:

1) **Visual features (Webcam data featuring students)**
   - Direction of the gaze
   - Frequency of eye blinks
   - Emotional expression (happy, confused, bored) [3],[34]
   - Head position
2) **Behavioral Features**
   - Number of clicks on instructional web pages/minutes of activity
   - Time spent with learning materials
   - Scroll activity
   - Frequency of switching between pages
3) **Performance-Based Features**
   - Response time to questionnaires
   - Correct answers
   - Number of attempts made on an assignment
   - Rate of completion on assignments [22]
4) **Interaction Features**
   - Number of discussion posts made
   - Average length of messages sent
   - Frequency of participation [8].

Once the features have been extracted, the resulting data set has now been converted into a feature matrix that can be used as input for modeling purposes.

Step 4: Data Normalization

Due to the possibility of features having disparate measurement scales, normalization is required so that each feature contributes equally toward model learning [11].

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Where:

- X: Original feature value
- X_min: Minimum value
- X_max: Maximum value
- X_norm: Normalized value

By applying normalization it is possible to obtain similar contributions from different features and boost the performance of many types of deep learning models, i.e., LSTM [25].

Step 5: Model Input Preparation

Following data normalization, the data is then edited into a format to be inputted into the engagement classification model [7] [26]. To accomplish this, the data needs to be put into a sequential format, e.g.:

$$X = (x_1, x_2, x_3, \dots, x_t) \tag{2}$$

Where x_t is the feature vector at time step t.

The sequential format that is generated will allow the LSTM networks to identify temporal patterns in user engagement behaviour [6].

The processed dataset will then be separated into three subsets of data:

1) The training dataset will take up 70% of the dataset
2) The validation dataset will take up 15% of the dataset
3) The testing dataset will take up 15% of the dataset

This distribution of the dataset provides a good balance between training, tuning and validation of the classification model, and will allow for better generalimprovement to unseen data [30].

V. Classification Models

The classification models in the proposed system classify and identify student engagement levels via learner data created from people logging onto an online learning environment (e.g. LMS). The classification models will review these features of student engagement and predict either a high level, medium level, or low level of student engagement during an LMS session.

To assess the performance of those models, the research utilized two machine learning models and one deep learning model. These include:

- Support Vector Machine (SVM)
- Random Forest

- Long Short-Term Memory (LSTM)

All three classification algorithms have unique characteristics and strengths when predicting student engagement behaviour patterns.

A. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised machine learning algorithm used in classification and pattern recognition [10].

Support Vector Machine (SVM) uses features extracted from detected engagement to identify classifications.

How SVM Works

SVM identifies a hyperplane that classifies the data:

$$w \cdot x + b = 0 \tag{3}$$

The ideal hyperplane provides the maximum separation between classes.

Benefits of SVM

- Can be applied to high-dimensional data
- Usually performs better with small datasets
- Usually produces good class separation

Limitations of SVM

- Not designed to be used with time series or sequential data [6]
- Very sensitive to the choice of kernel function
- Can be a computationally intensive process with larger datasets

B. Random Forest

Random Forest is an ensemble learning method which using multiple decision trees [10], [15].

Working Steps

1) Generating random subsets (bootstrapping)
2) Train decision trees on subsets
3) Each tree predicts output
4) Final prediction via majority voting

Example Logic

- Engaged if interaction > threshold AND response time < threshold
- Disengaged if interaction < threshold AND response time > threshold

Benefits

- High accuracy
- Robust to noise and outliers
- Handles non-linear data
- Reduces overfitting

Drawbacks

- Cannot model sequential data [6]
- Complex for large datasets
- Less interpretable

C. LSTM Model (Deep Learning)

LSTM is a type of Recurrent Neural Network (RNN) designed for sequential data.

It captures temporal patterns in engagement such as changes in student focus over time.

LSTM Characteristics

- Learns long-term dependencies
- Suitable for time-series data
- Uses sequential input
- Reduces vanishing gradient problem

D. LSTM Architecture

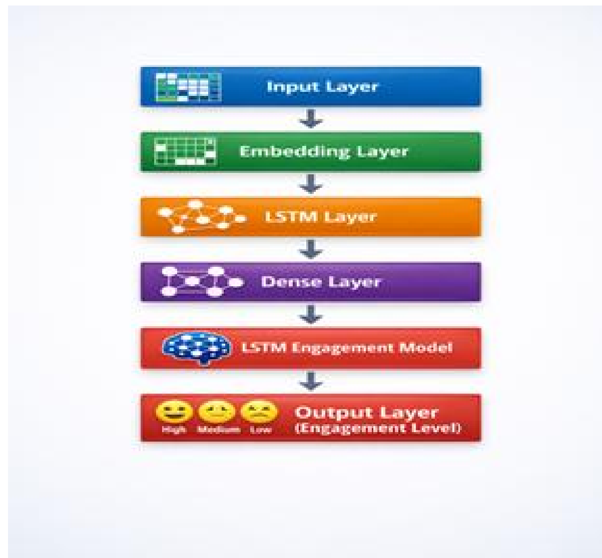

Fig. 2. Proposed System Architecture

Input Layer Processes sequential feature vectors including:

- Interaction frequency
- Response time
- Facial emotions
- Keyboard activity
- Page navigation

Embedding Layer

- Converts categorical data to dense vectors
- Reduces dimensionality
- Improves feature representation

LSTM Layer

Processes sequence:

$$h_t = f(h_{t-1}, x_t) \tag{4}$$

Uses gates:

- Forget gate
- Input gate
- Output gate

Dense Layer

- Fully connected layer

- Uses ReLU activation

Output Layer

Uses Softmax for classification:

$$P(y=i)=\frac{e^{z_i}}{\sum e^{z_j}} \quad (5)$$

Output Classes

- Highly Engaged
- Moderately Engaged
- Disengaged

E. Model Comparison

Support Vector Machines are effective for structured classification but cannot capture sequential engagement patterns.

Random Forest improves accuracy using multiple trees but lacks time-based analysis.

LSTM models outperform traditional methods by capturing temporal dependencies and dynamic engagement behaviour.

Studies show that LSTM achieves higher accuracy due to its ability to learn from sequential multimodal data [27], [30].

Model	Strength	Limitation
SVM	Good for structured data	Cannot model time-dependent data
Random Forest	High accuracy and robust	Poor for sequential analysis
LSTM	Captures time-based patterns	Requires high computation

TABLE V
COMPARISON OF CLASSIFICATION MODELS

VI. EXPERIMENTAL RESULTS AND DISCUSSION

This section describes the evaluation process of the proposed method for detecting student engagement in online learning environments. The purpose of this experiment was to compare and contrast several machine-learning algorithms and classify students' level of engagement accurately.

Using data collected from logs of student interaction behaviours (e.g., click-stream logs), along with their visual engagement behaviours, a sample dataset was assembled for this study. The sample dataset was then divided into three subdatasets.

A. Dataset Division

The dataset was split into three parts:

- **Training Dataset (70%):** Used to train the model.
- **Validation Dataset (15%):** Used to tune model parameters.
- **Test Dataset (15%):** Used to evaluate performance.

This division helps in improving model generalization.

B. Model Evaluation Metrics

Accuracy

Accuracy is defined as:

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Where:

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

Higher accuracy indicates better model performance.

C. Experimental Results

Using the test dataset, the three engagement detection models were evaluated. The outcomes from the testing trials are summarized in the table below:

Model	Accuracy
Random Forest	86%
SVM	82%
LSTM	93%

TABLE VI
MODEL ACCURACY COMPARISON

Result Interpretation

- Random Forest achieved 86% accuracy and performed well for classification tasks.
- SVM achieved 82% accuracy but struggled with complex patterns.
- LSTM achieved the highest accuracy of 93%, showing superior performance.

The higher accuracy of LSTM is due to its ability to capture temporal dependencies in sequential data.

D. Graphical Comparison of Model Accuracy

The classifiers compared are:

- RF = Random Forest
- SVM = Support Vector Machine
- LSTM = Long Short-Term Memory

LSTM shows significantly better accuracy compared to other models.

E. Discussion of Results

Performance of Random Forest

Random Forest performs well due to ensemble learning and robustness to noise. However, it cannot capture sequential dependencies.

Performance of SVM

SVM is effective for structured data but struggles with large-scale sequential datasets.

Performance of LSTM

LSTM provides superior performance due to:

- Ability to capture temporal dependencies
- Handling sequential data
- Learning long-term patterns

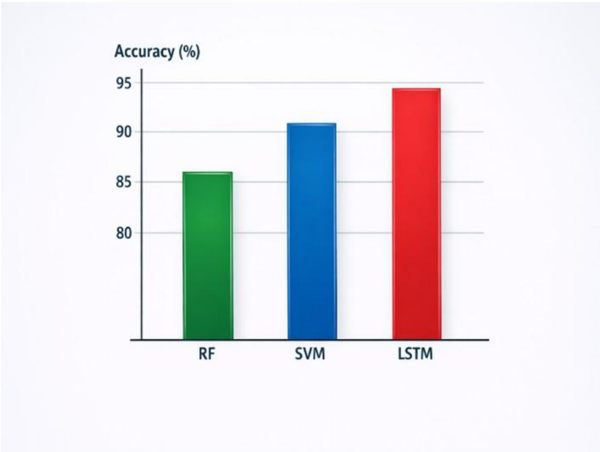

Fig. 3. Graphical Representation

F. Key Findings

1) Deep learning models outperform traditional machine learning methods.
2) Temporal features are critical for engagement detection.
3) Multimodal data improves prediction accuracy.
4) LSTM provides the best performance for engagement detection.

VII. Applications

The AI-driven solution is a flexible way to gather data on how students are participating in the current digital environment. When analyzing multiple modes of how students learn and their engagement trends, this provides educators, institutions, and other organizations a way to help improve learning outcomes.

A. Smart Classrooms

Smart Classrooms use interactive boards, Learning Management Systems (LMS), and AI-based analytics tools to monitor student engagement.

Instructors can:

- Identify students who are not engaged or are distracted
- Analyze the patterns of a classroom's participation
- Provide individualized instruction for each student
- Change teacher's methods of delivery based on the level of student engagement

Real-time analytics improve teaching quality and student performance.

B. Online Learning Platforms

Automated engagement monitoring benefit students taking courses using MOOC, VIRTUAL CLASSROOMS, OR other e-learning platforms. [20],[22].

The system can:

- Tracking student access and interaction to course content [1],[13]
- Evaluating attention level during video lectures [3]
- Identifying potential 'at-risk' students [17]
- Providing students with automated feedback and study recommendations [27],[30]

This is especially beneficial for large-scale online courses.

C. Organization Corporate Training Programs

Organizations use online platforms to enhance employee skills efficiently.

Ways to Use Online Learning:

- Tracking employees who participate in online learning and training.
- Measuring employee focus during virtual training sessions.
- Identifying training areas that do not add value.
- Improving learning outcomes through virtual business training.
- Increasing efficiency and productivity through online training.

D. Educational Analytics Systems

Educational analytics systems help track and evaluate student behavior.

We Can Be Achieved Through Educational Analytics:

- Engagement trend analysis.
- Insight into how students learn and perform.
- Understanding performance of highly engaged students.
- Identifying students at risk of failing or dropping out.

E. Personalized Learning Systems

Learning content is tailored according to individual student needs.

AI-Driven Personalized Learning Management System:

- Recommendation of additional learning resources.
- Adjustment of lesson difficulty based on student level.
- Suggestion of study skills.
- Providing adaptive and timely feedback.
- Enhancing motivation, retention, and learning efficiency.

VIII. Challenges and Future Work

The future implementation of engagement detection technologies is at risk due to privacy and ethical issues, the requirement for large amounts of labeled data, and the possibility of low-quality labeled data.

A. Privacy and Ethical Concerns

Privacy concerns arise from the collection of sensitive information such as webcam feeds, facial recognition data, and behavioral logs by AI-based engagement detection systems.

Important issues include:

- Unauthorized access to or breaches of sensitive information
- Misuse of facial recognition data for unauthorized purposes
- Lack of transparency in engagement detection methods

Future systems must ensure:

- Respect for individual rights, with data used only with user consent

- Secure mechanisms for storing sensitive information
- Compliance with privacy laws and regulations (e.g., GDPR)

To ensure ethical deployment, developers must address these concerns carefully.

B. Requirement for Large Labeled Datasets

Deep learning models such as LSTM require large labeled datasets for high predictive accuracy. However, collecting and labeling engagement data is time-consuming and costly.

Challenges include:

- Difficulty in accurately labeling engagement levels
- Variability of engagement across different contexts
- Limited availability of public datasets

Future approaches may include:

- Semi-supervised learning
- Transfer learning techniques

These approaches can help reduce reliance on large labeled datasets.

C. Difficulty Detecting Sarcasm or Hidden Emotions

Detecting sarcasm or concealed emotions remains a significant challenge. For example:

- A student may appear attentive but may actually be disengaged
- Facial expressions may not accurately represent true emotions

This highlights the need for more advanced emotion recognition techniques.

D. Future Research Directions

1) **Transformer-Based Models:** Recent advancements in transformer-based architectures have shown improved performance over traditional RNN and LSTM models. These models can be explored for more reliable sequence analysis.
2) **Multimodal Deep Learning:** Future systems can integrate multiple data sources such as:
   - Video
   - Audio
   - Physiological sensors (heart rate, EEG)
   - Text data

   This integration can significantly improve engagement detection accuracy.
3) **Real-Time Engagement Monitoring:** Real-time systems can provide:
   - Immediate feedback to instructors
   - Timely interventions
   - Improved student engagement and performance

IX. Conclusion

X. Conclusion

The purpose of this research involved creating an artificial intelligence (AI)-based study assistant to track and analyze how students engage in a digital learning environment. Given the rapid increase of online educational platforms, educators around the world are increasingly concerned with monitoring and evaluating student engagement [1][2].

The proposed model utilizes learning analytics along with machine learning (ML) and deep learning (DL) techniques to analyze student interactions with course materials and determine their level of engagement. Various types of engagement-related data such as behavioral logs, facial expressions, and response patterns are collected and processed through modules including data preprocessing, feature extraction, and classification [3][4].

Among the evaluated models, the Long Short-Term Memory (LSTM) model achieved the best performance with an accuracy rate of 93%, significantly outperforming traditional machine learning approaches [5][6]. This is due to its ability to effectively capture temporal dependencies in sequential data.

The proposed study assistant provides several benefits in modern educational environments. It enables educators to identify at-risk or disengaged students, deliver personalized feedback, and implement data-driven instructional strategies. This ultimately leads to improved academic performance, increased student motivation, and reduced dropout rates in online learning systems [7][8].

Future research may focus on the development of transformer-based models, integration of multimodal data sources, and real-time engagement monitoring systems. These advancements will contribute to more intelligent and effective educational technologies, enhancing overall digital learning experiences [9][10].

XI. References

References

[1] R. S. Baker and K. Yacef, “The state of educational data mining in 2009,” *Journal of Educational Data Mining*, vol. 1, no. 1, pp. 3–17, 2009.

[2] S. D’Mello and A. Graesser, “AutoTutor and affective autotutor,” *ACM Transactions on Interactive Intelligent Systems*, vol. 2, no. 4, 2012.

[3] J. Whitehill *et al.*, “Faces of engagement: Automatic recognition of student engagement from facial expressions,” *IEEE Transactions on Affective Computing*, vol. 5, no. 1, pp. 86–98, 2014.

[4] Y. Bengio, I. Goodfellow, and A. Courville, *Deep Learning*. MIT Press, 2016.

[5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

[6] T. Mikolov *et al.*, “Recurrent neural network based language model,” in *Proc. INTERSPEECH*, 2010.

[7] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[8] J. Devlin *et al.*, “BERT: Pre-training of deep bidirectional transformers,” in *Proc. NAACL*, 2019.

[9] A. Krizhevsky *et al.*, “ImageNet classification with deep convolutional neural networks,” in *Proc. NIPS*, 2012.

[10] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer, 2009.

[11] C. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[12] H. Peng *et al.*, "Multimodal engagement detection in online learning," *IEEE Access*, 2020.
[13] L. Zeng *et al.*, "Learning analytics in higher education," *Educational Technology & Society*, 2018.
[14] P. Dillenbourg, *Virtual Learning Environments*. Springer, 2015.
[15] J. Han *et al.*, *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2011.
[16] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson, 2016.
[17] R. S. Baker, "Educational data mining and learning analytics," in *Handbook of Learning Analytics*, 2017.
[18] B. Woolf, *Building Intelligent Interactive Tutors*. Morgan Kaufmann, 2010.
[19] M. Worsley and P. Blikstein, "Multimodal learning analytics," in *Proc. LAK Conference*, 2014.
[20] G. Siemens and R. Baker, "Learning analytics and educational data mining," in *Proc. LAK*, 2012.
[21] A. Kapoor *et al.*, "Predicting student engagement," in *Proc. ACM CHI*, 2015.
[22] K. Koedinger *et al.*, "Learning analytics and data mining," *Educational Psychologist*, 2015.
[23] D. Romero and S. Ventura, "Educational data mining: A review," *IEEE Transactions on Systems*, 2013.
[24] S. K. D'Mello, "A selective meta-analysis on emotion in learning," *Educational Psychologist*, 2013.
[25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
[26] A. Graves, *Supervised Sequence Labelling with RNNs*. Springer, 2012.
[27] Z. Zhang *et al.*, "Deep learning for engagement detection," *IEEE Access*, 2019.
[28] M. Chaudhuri *et al.*, "Student engagement detection in e-learning," *Computers & Education*, 2020.
[29] L. Chen *et al.*, "Multimodal engagement recognition," *IEEE Transactions on Multimedia*, 2018.
[30] H. Liu *et al.*, "Deep neural networks for learning analytics," *IEEE Access*, 2021.
[31] J. Brown *et al.*, "AI in education," *Educational Technology Review*, 2019.
[32] P. Blikstein, "Multimodal learning analytics," in *Proc. LAK*, 2017.
[33] S. Buckingham Shum and R. Ferguson, "Social learning analytics," *Educational Technology*, 2012.
[34] A. Kapoor and R. Picard, "Multimodal affect recognition," in *Proc. ACM Multimedia*, 2005.
[35] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. KDD*, 2016.