J. Electrical Systems 20-3s (2024): 401-418

[1]Guangheng Tang

Sequence Modeling with Recurrent Neural Networks (RNNs) for Student Learning Behavior Pattern Recognition in a Flipped Classroom

Abstract: - The flipped classroom model has become increasingly popular in education, altering the traditional methods of teaching. It is important to understand and acknowledge how students learn within this framework in order to optimize instructional strategies and promote personalized learning. This study investigates the use of Sequence Modeling with Recurrent Neural Networks (RNNs) to identify patterns in student learning behavior within a flipped classroom setting. The proposed deep learning architecture utilizes RNNs to analyze sequential patterns in students' interactions with the flipped classroom materials, while also incorporating attention mechanisms to better detect important patterns and temporal dynamics in the learning process. Multimodal learning techniques are also employed, combining data from various sources to gain a comprehensive understanding of student behavior. Additionally, clustering techniques using autoencoders are explored to group students with similar learning behaviors. Predictive models, such as RNN or LSTM networks, are developed to forecast future learning behaviors and provide insights into potential challenges or successes for individual students. The effectiveness of this framework is evaluated using real-world data from flipped classroom implementations, with performance metrics like recall, precision, and accuracy used to assess the success of the sequence modeling approach in recognizing and predicting student behavior patterns. Overall, the application of deep learning methods, specifically sequence modeling with RNNs, demonstrates potential for improving personalized learning experiences and facilitating proactive interventions to support diverse student needs.

Keywords: Flipped Classroom Model, Recurrent Neural Networks (RNNs), Sequence Modeling, Long Short-Term Memory (LSTM) Networks and Deep Learning Architecture

I. INTRODUCTION:

A crucial aspect of modern education is the classroom instruction. Improving the quality of student learning within the classroom not only boosts their academic success but also encourages the use of more engaging teaching approaches. In both online and large classroom settings, it can be challenging for teachers to accurately assess students' listening and ensure adherence to classroom rules using traditional methods [1]. By incorporating intelligent technology to detect and provide feedback on students' learning behavior and listening quality, teachers can better monitor their students in real-time. However, implementing such a system can be challenging due to complex models or high costs. Therefore, this study utilizes a DL algorithm as foundation and prunes it to maintain satisfactory performance while ensuring practical implementation.

Recognizing emotions in real-time scenarios poses challenges, demanding sophisticated algorithms to precisely interpret facial expressions. Despite these hurdles, progress in computer vision technology has enabled the development of more effective FER systems. These systems can adjust course material based on a learner's expressions, leading to improved learning and engagement [2]. This allows instructors to create personalized learning experiences that meet the individual preferences and needs of each student. Our research indicates a significant link between a learner's emotions and their engagement and interest levels during online lectures. This insight can be leveraged to customize content and boost participation on digital learning platforms. In essence, our study underscores the promise of employing facial expression recognition to enhance the efficacy of adaptive learning systems and foster academic achievements.

Technology plays crucial role in delivering and improving education in the realm of online learning. The most effective tool for teachers to facilitate collaboration and boost student engagement is video conferencing. It not only serves as an efficient means of teaching, but also as a mode of communication in this virtual setting. There are several platforms available for video conferencing, such as WebEx, Go To Meeting, Skype, Zoom, Microsoft Teams, for business. These platforms enable real-time communication through live audio and video between teachers and learners [3]. This immediate connectivity fosters a sense of human connection, allowing educators and students to establish their presence in the online learning environment. By effectively utilizing video conferencing, teachers can provide prompt feedback to students, bridging the psychological and communication

[1] Hubei Polytechnic Institute, Xiaogan, Hubei, China, 432100

*Corresponding author's e-mail: tangguangheng1122@126.com

gap between them. Additionally, video conferencing offers features like wireless screen sharing, whiteboard sharing, interactive chat rooms, opinion polls, and discussion platforms to enhance the learning experience.

FCs has become a popular subject, especially in higher education, and there has been a significant rise in research on their effectiveness. These classrooms differ from traditional ones in that the material typically taught during class is instead given beforehand, outside of class. As a result, students primarily learn on their own time, allowing for more interactive and engaging learning experiences during face-to-face sessions [4]. The success of FCs heavily relies on educational technology, which provides easier access to resources. This approach prioritizes student involvement as they actively construct their own knowledge.

Flipped classrooms (FC) offer an enhanced learning experience by providing more flexible learning opportunities and promoting active learning, resulting in increased student engagement. Studies have shown that student performance is generally better in FCs compared to traditional classrooms. However, the findings from research on FCs are mixed, with some reporting positive results ranging from small to moderate effects, while others finding no significant difference. Most of these studies focused on STEM subjects, with business being unspecified [5]. In cases where business courses were included, no discernible impact was found. Additionally, research suggests that FCs may be more beneficial for post-graduate students compared to undergraduate students who tend to have lower motivation levels. Student perceptions of FCs have also been examined, and most students believe that they provide a more positive learning experience and facilitate better retention of knowledge. However, it should be noted that these benefits are mostly observed in students who already possess high levels of motivation and may be reduced if negative perceptions of FCs exist among students [6].

The implementation of FCs is made simpler by advancements in technology. E-learning resources play a crucial role in overcoming the limitations of traditional classrooms. Specifically, these technological tools aid in pre-class preparation through the utilization of e-books, audio slides, and videos. Research has shown that, compared to reading materials, videos have a positive impact on student performance, although this effect is less significant for undergraduate students. Therefore, it is important for videos to promote student engagement rather than just passive viewing [7]. To enhance interactivity, tests and quizzes can be incorporated into online materials to make them more meaningful and beneficial for students

The use of advanced technologies has become crucial in the ever-changing field of education to improve the learning process. A significant tool in this regard is Sequence Modeling with RNNs, which holds great promise in analyzing and optimizing student learning patterns in a flipped classroom environment. The flipped classroom model shifts the traditional teaching method by providing online instructional materials, giving students the freedom to learn at their own pace outside of class. In this progressive educational setting, monitoring and comprehending student behavior is vital for customizing the learning experience to meet individual needs [8]. This is where Sequence Modeling with RNNs emerges as an innovative solution.

RNNs are a specialized category of artificial neural networks crafted to process data arranged in a sequential fashion. This makes them particularly well-equipped for identifying and understanding the time-based connections that are inherent in the learning process. By examining the sequence of interactions, RNNs can uncover patterns in how students interact with pre-class materials, engage in discussions, and grasp important concepts during class activities. This brings together the worlds of RNNs and FC methodologies, demonstrating the potential for these technologies to transform how we recognize patterns in student learning behavior. As we delve deeper into the intricacies of sequence modeling, we will discover how RNNs can offer valuable insights into student participation, understanding, and overall academic performance. Ultimately, this has the potential to enhance the educational experience by providing personalized and effective learning paths for students. These advancements pave the way for a data-driven approach to education, creating an environment where educators can refine their teaching strategies and students can benefit from tailored and adaptive learning methods. The major contribution of the article is described below.

1.1 Research Objectives:

- RNN-based sequence modeling is a valuable tool for identifying and recognizing unique learning behavior patterns. This allows for the creation of personalized learning paths that cater to the specific needs, preferences, and progress of students.

- With RNNs, it is possible to design adaptive instruction that dynamically adjusts the delivery and sequence of educational content, ensuring that students receive information in a way that suits their preferred learning styles and pace. By adapting to individual learning behaviors, this system promotes increased student engagement.
- Furthermore, tailored content and personalized learning experiences make the educational journey more meaningful and enjoyable for each student. The use of RNNs for sequence modeling also provides educators and administrators with data-driven insights into student learning behaviors, which can inform decision-making and lead to improvements in curriculum design, teaching methods, and overall educational strategies.
- In a flipped classroom framework, wherein students engage with instructional materials outside conventional class hours, RNN-based sequence modeling boosts the efficacy of both pre class and in class activities by tailoring them to individual learning styles.

II. RELATED WORKS

In [9], facial emotion detection system using ensemble models to monitor student attentiveness in online learning, achieving recognition rates above 90%. By providing continuous feedback, instructors can adapt teaching methods to enhance student engagement and learning outcomes. In [10], discovering a positive relationship among the quality of educational systems, ease of use, and overall satisfaction, while questioning the influence of information quality, is a key focus. This underscores the significance of refining educational materials and taking into account diverse factors to enhance student satisfaction and engagement in online learning settings. These findings provide valuable insights for educators and platform developers.. In [11], findings suggest a preference for anonymity despite its potential to create a less productive atmosphere, prompting suggestions for optimizing anonymity features in collaborative learning tools. In [12], investigates how resourceful behaviors like financial bootstrapping and bricolage influence innovative behavior among student entrepreneurs in Spain. Findings suggest that bricolage mediates the relationship between financial bootstrapping and innovation, highlighting implications for resource management in fostering innovation among student ventures. In [13], the impact of mindful agency on college students' connection with online teaching, mediated by online learning self-efficacy and self-disciplined learning online. The findings contribute to both theoretical comprehension and practical advice for enhancing identification with online teaching.. In [14], an AI-enabled tool for generating reports on engaging teaching videos in higher education, utilizing deep learning for video analysis. The tool aids in identifying and enhancing engagement behaviors of teachers during video conferencing, offering insights for education institutes and instructors to improve online learning experiences.

In [15], factors influencing undergraduate students' adoption of AI chatbots for educational use, finding that perceived benefits, compatibility, trialability, and trust significantly influence adoption intention. Surprising results suggest the need for further investigation into the dynamics of perceived ease of use and usefulness in chatbot advocacy. In [16], the relationship between students' environmental mindset and their entrepreneurial actions, revealing that entrepreneurship education approaches can influence this connection, particularly with competency-focused methods reinforcing it. Moreover, the moderating impact of competency teaching methods is augmented by mastery-learning motivations, advocating for active learning strategies in entrepreneurship education. In [17], identifies two distinct learning strategies in a flipped classroom: "Prepared" students access resources ahead of class, while "Assessment-focused" students engage post-class for upcoming assessments. "Prepared" students show higher performance, suggesting structuring flipped classrooms with frequent, low-stakes assessments can mitigate procrastination and enhance effectiveness. In [18], developed a MOOC data from AdelaideX in 2019 and 2020, focusing on engagement, semantics, and sentiment/stress to understand their impact on student outcomes is concentrated. Results show that stress had small influence on academic attainment and remained proportional consistent across online courses in the studied period. Two student cases provide additional context to the findings. In [19], the ANN to identify key factors impacting undergraduate academic performance, highlighting sleep quality, questioning, class attendance during lectures as highly correlated with high grades. The research showcases the effectiveness of student data gathering and ML for understanding the nuanced relationship between study behaviors and academic achievement.

In [20], introduced ODL-BCI, an optimized deep learning model designed for real-time classification of students' confusion levels using EEG data. It surpasses current state-of-the-art methods, achieving an accuracy improvement of 4% to 9%. A source codes for ODL-BCI implementation are available for access, providing a valuable tool for BCI research in educational condition. In [21], influence of various peer assessment modes on students' behaviors and performance in musical theater. Result show that the integrated peer scoring and commenting mode was most effective which leading to improved performance, critical feedback, and engagement with assessment materials. In [22], an CM-TTG approach demonstrates a notable enhancement in ninth-grade students' learning achievements when compared to the conventional two tier test based gaming method. However, no significant variances were observed in learning motivation, flow experience, or cognitive load between the two groups. Interestingly, distinct learning behavior patterns were noted among high and low achieving students within CM-TTG approach. In [23], this study examines SRL behaviors of undergraduate students on a Moodle platform, finding a recursive SRL cycle with distinct patterns for high- and low-performing students. High-performing students adjusted and enhanced their SRL behaviors more significantly following a formative exam, offering insights for practitioners on fostering effective SRL through feedback. The summary of this earlier researches are described in table 1.

Table 1 – Previous Research Summary

Ref. No	Algorithm	Methodology	Advantages	Disadvantages	Performance	Efficiency	Accuracy	Features Used	Measurements
[9]	Ensemble	Facial emotion detection system in online learning	Enhances student engagement and learning outcomes	Difficulty in optimizing anonymity features in collaborative learning tools	Above 90%	High	High	Facial emotion recognition	Student attentiveness
[10]	Ensemble	the relationship between the quality of the educational system and overall satisfaction of users in e-learning.	Insights for enhancing educational materials and improving student satisfaction	Challenging impact of information quality	NIL	High	N/A	Educational system quality, ease of use, satisfaction	Satisfaction, engagement
[11]	Ensemble	Research on anonymity and its effects on collaborative learning	Preference for anonymity despite productivity concerns	Challenges in optimizing anonymity features for productivity	Highlights potential benefits	N/A	N/A	Anonymity features	Productivity, engagement
[12]	Ensemble	Exploration of resourceful behaviors in student entrepreneurship	Implications for resource management in fostering innovation among student ventures	Limited focus on specific emotionsStudy on how mindful agency influences	High	N/A	N/A	Financial bootstrapping, bricolage, innovative behavior	Innovation, resourcefulness

				students' identification with online teaching.					
[13]	Ensemble	Study on how mindful agency impacts students' identification with online teaching.	Theoretical understanding and guidance for enhancing identification with online teaching	Limited impact on cognitive load, motivation, or flow experience	NIL	High	NIL	Mindful agency, online learning self-efficacy, self-disciplined learning online	Identification, motivation, learning outcomes
[14]	Deep Learning	AI-enabled tool for generating reports on engaging teaching videos in higher education	Identification and enhancement of engagement behaviors in video conferencing	Limited exploration	Improved precision and accuracy	NIL	NIL	Deep learning for video analysis	Engagement, teaching quality
[15]	Deep Learning	Factors influencing undergraduate students' adoption of AI chatbots for educational use	Insights into factors affecting adoption intention	Need for further investigation into dynamics of perceived usefulness and ease of use	NIL	NIL	NIL	Perceived benefits, compatibility, trialability, trust	Adoption intention, effectiveness
[16]	Deep Learning	Study on entrepreneurial actions	Implications for entrepreneurship education methods in moderating relationship	Limited exploration	Enhanced accuracy	NIL	NIL	Environmental mindset, entrepreneurship education methods	Entrepreneurial actions, mindset
[17]	Deep Learning	Examination of learning strategies in a flipped classroom	Structuring flipped classrooms with frequent, low-stakes assessments for enhanced effectiveness	Mitigation of procrastination through structured assessment	High accuracy	NIL	NIL	Engagement, assessment strategies	Learning strategies, effectiveness
[18]	Deep Learning	Exploration of MOOC data and its impact on student	Insight into engagement, semantics, and	Sentiment/ stress had little influence	NIL	NIL	NIL	Engagement, semantics, sentiment/s	Student outcomes,

		outcomes	sentiment/stress on student outcomes	on academic performance				tress	sentiment
[19]	ANN	Identification of key factors impacting undergraduate academic performance	Controlled student data collection and machine learning for understanding study behaviors and achievement	High correlation between class attendance, sleep quality, and high grades	New framework for understanding	NIL	NIL	Class attendance, sleep quality, questioning during lectures	Academic performance, study behaviors
[20]	DL	ODL-BCI, an optimized deep learning model designed for the real-time classification of students' confusion levels using EEG data.	Achieves superior performance compared to state-of-the-art methods, with an accuracy increase ranging from 4% to 9%.	Limited exploration	4%-9% higher	NIL	NIL	EEG data for confusion levels	Real-time classification of confusion levels
[21]	Deep Learning	how different peer assessment modes influence behaviors and performance in musical theater.	Improved performance, critical feedback, engagement with peer scoring and commenting mode	No significant differences in learning motivation, flow experience, or cognitive load between groups	NIL	NIL	NIL	Peer scoring, peer commenting	Musical theater performance, engagement
[22]	Deep Learning	CM-TTG approach for ninth-grade students' learning achievement	Improved learning achievement compared to conventional approach	No significant differences in motivation, flow experience, or cognitive load between groups	NIL	NIL	NIL	Two-tier test-based gaming, concept mapping	Learning achievement, engagement, cognitive load
[23]	Deep Learning	Study on SRL behaviors of undergraduate students on a	Recursive SRL cycle observed with distinct	High-performing students adjusted	NIL	NIL	NIL	SRL behaviors on Moodle platform	Learning achieveme

		Moodle platform	patterns for high- and low-performing students	SRL behaviors significantly following a formative exam					nt, SRL behaviors

III. FUNDAMENTALS

3.1 Flipped Classroom Model:

The concept of a flipped classroom involves changing the traditional learning environment. In a regular classroom, students are taught new ideas directly by their teacher and then practice and use them as homework. However, in a flipped classroom, students are exposed to new material through online resources before or outside of class. In the classroom, they participate in interactive activities such as discussions and hands-on projects to apply their knowledge. Figure 1 as explained the Flipped Classroom Model as drawn.

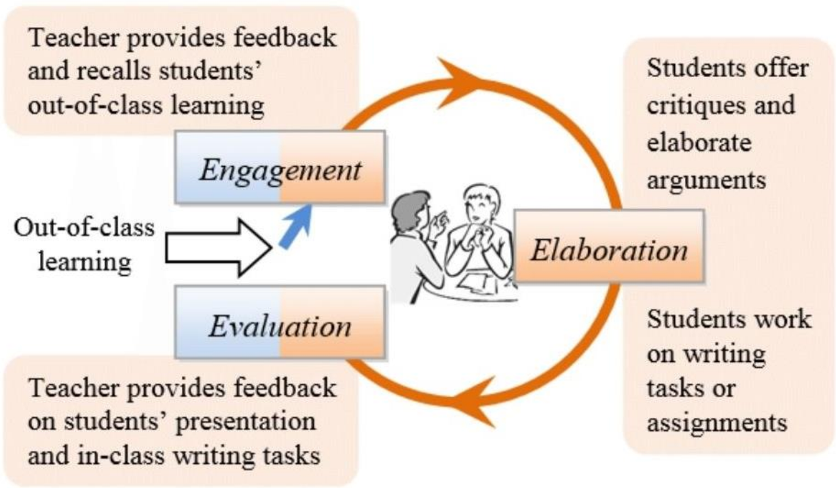

Figure 1 - Flipped Classroom Model

The FCM consists of two main components: Pre Class Work, which is done at home, and In-Class Work. During the Pre Class Work, students independently prepare for the upcoming class by watching pre-recorded lectures, videos, or reading assigned materials. They also take notes, answer questions, and gain a basic understanding of the new concepts. In the classroom, students engage in active learning activities such as group work, problem solving exercises, discussions and hands-on activities. This allows them to reinforce and deepen their understanding of the material. Additionally, they have the opportunity to ask questions and apply what they learned at home. The Flipped Classroom Model offers several advantages including individualized pacing for students to speed that is comfortable for you and increased engagement due to active learning. It also promotes a deeper understanding of concepts through immediate application in class and allows teachers to provide personalized feedback and support since they have more time to interact with students.

3.2 Recurrent Neural Networks:

RNNs are a specialized form of ANN intended for managing consecutive data and time-based tasks. Unlike conventional feedforward neural networks that treat each input separately, RNNs have interconnected links that create a circular pattern, enabling them to demonstrate dynamic temporal patterns. This feature makes RNNs highly suitable for tasks involving sequential input and/or output, such as language processing, voice recognition, prediction of time series data, and other related tasks. Figure 2 and Table 2 are discussed about aspect of RNN.

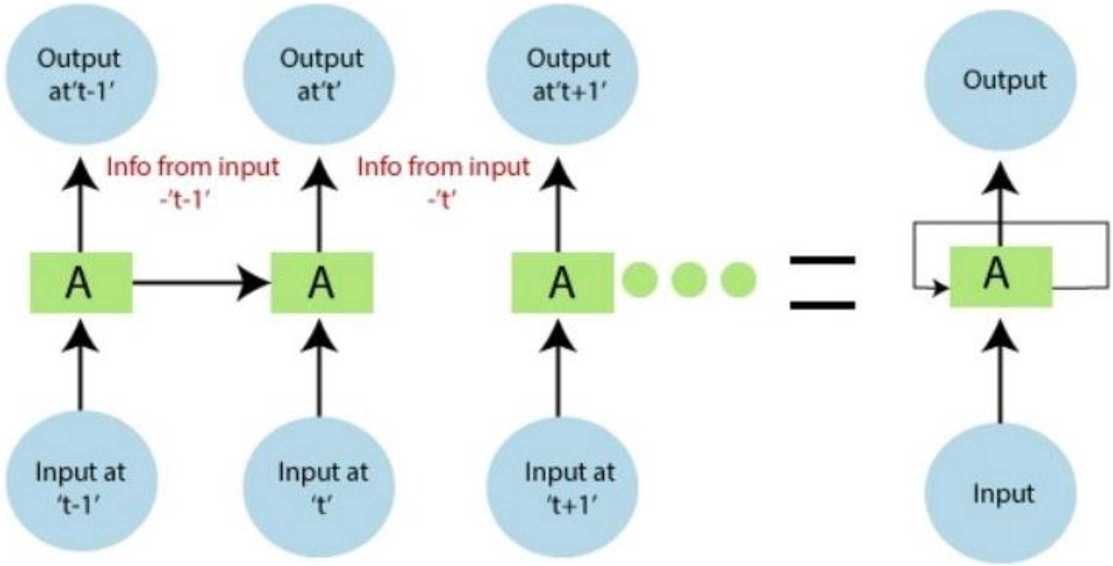

Figure 2 - RNNs Structure

RNNs can generate descriptions or captions for images based on the visual features extracted from convolutional neural networks. RNN Structure and Recurrent Connections: An important aspect of RNNs is their use of recurrent connections within the network. Table 3 explains the types of algorithm in RNNs. These connections form loops that allow information to persist, capturing relationships between time steps.

Table 2 - Aspects of RNN

Aspect	Description
Time Unfolding	Unfolding the RNN over time steps helps understand its processing of sequences. Each time step corresponds to an input element, allowing visualization of how the network handles sequential data.
Hidden State	The RNN retains a hidden state \(h_t\) at every time step \(t\). This hidden state encodes information from earlier inputs, serving as the network's memory. It summarizes the data seen up to that point and helps in making predictions for the next step in the sequence.
RNN Training: Backpropagation Through Time (BPTT)	Training RNNs involves Back propagation Through Time (BPTT), an extension of back propagation used in feed forward neural networks. BPTT calculates gradients of the loss function for each time step. These gradients are then used to update the network's weight, allowing it to learn from consecutive data.
RNN Training: Exploding Gradient Problem and Vanishing	RNNs may encounter challenges such as the vanishing gradient problem, where gradients diminish significantly as they propagate through time, or the exploding gradient problem, where gradients become excessively large. To address these issues, strategies like gradient clipping, utilization of various activation functions, and the adoption of specialized architectures such as LSTM or GRU are employed.

Table 3 - Types of RNNs

Types	Description
Vanilla RNN	The fundamental structure of an RNN is similar to what was previously explained. It includes connections that enable data to pass through different time points. However, its effectiveness in capturing distant relationships is hindered by the issue of vanishing gradients.
Long Short-Term Memory	An advanced type of RNN has been devised to address the limitations of basic RNNs, particularly in capturing long-term dependencies. LSTMs feature a more intricate architecture, incorporating elements such as input, forget, and output gates. These

	gates control the movement of data within the cell, enabling it to retain and recall information over long sequences. The memory cell in LSTMs is responsible for storing important data, avoiding the issue of vanishing gradients.
Gated Recurrent Unit (GRU)	GRUs were created with the intention of improving upon the limitations of basic RNNs, much like LSTMs. They were specifically designed to tackle the issue of vanishing gradients and are more efficient in terms of computation compared to LSTMs. Additionally, GRUs include gates such as update and reset gates which allow them to better capture connections between time steps and retain information over longer sequences while also having a smaller number of parameters. As a result, they are easier to train and more efficient overall.

3.3 Sequence Modeling:

Sequence modeling is the process of forecasting the next element in a sequence of data points. It plays a crucial role in various fields, such as natural language processing, speech recognition, time series analysis, genomics, and others. The primary goal is to develop a model that can identify patterns and connections among the data sequence and use that knowledge to make accurate predictions follow up with Figure 3.

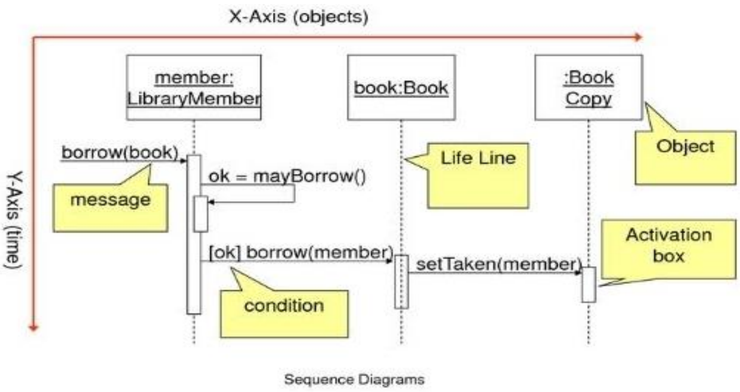

Figure 3 – Sequence Diagram

3.4 LSTM Networks:

LSTM networks, a form of RNN design, were developed to overcome the limitations of conventional RNNs in preserving and understanding long-term relationships in sequential information. As shown in Figure 4, these networks have proven to be very effective in tasks like predicting time series, processing natural language, recognizing speech, and others that involve capturing distant connections.

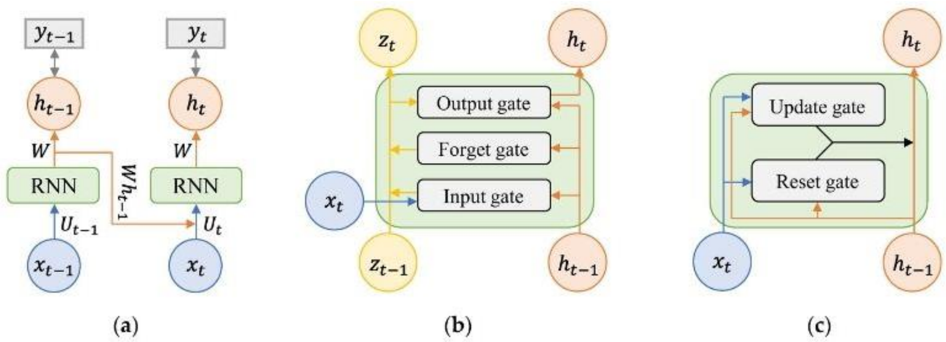

Figure 4 - LSTM Architecture

IV. PROPOSED SLBPR-RNN MODEL

4.1 Students Activities Findings

The analysis revealed two distinct groups of student behaviors. The first, called the "Prepared" cluster, consisted of 58% of the students. These students consistently utilized pre-eclampsia materials before both classes and assessments throughout the course. The second group, referred to as the "Assessment-focused" cluster, made up 42% of the students. They tended to access online resources after class sessions but before weekly assessments. Within the Prepared cluster, there was a consistent trend of accessing materials before class, although there was a noticeable difference between different sections of the course. In section 1, students accessed flipped resources 19 hours before class while in another section this occurred just 1 hour before. This difference in timing may be attributed to assessment schedules and the nature of online materials used. For example, when presented with video-based materials and small-scale weekly assessments, students prepared well in advance by gaining an initial understanding beforehand, applying it during class activities, and then revising any unclear points using online resources before assessments. However, their approach changed in a different section where they prioritized accessing written materials right before class to have them available during in-class activities instead of engaging with text-based content beforehand.

On the other hand, the Assessment-focused cluster typically accessed resources after relevant classes but before assessments for both sections of the flipped course. Notably, when the course structure shifted from weekly low-stakes assessments based on video materials to a higher-stakes assessment based on three weeks of written material, this group tended to procrastinate more before accessing online resources. This highlights how important course structure is in a flipped classroom setting, particularly in terms of assessment timing and weighting. Unlike the Prepared cluster, the Assessment-focused group did not prioritize engaging with and understanding online materials before class as preparation for assessments. These findings demonstrate how learning strategies can adapt to changes in course structure.

After identifying two learning strategies, the researchers analyzed their respective performances. The flipped class was divided into two sections, each worth 30% of the total grade. The study did not include the 40% group project. Students in the Prepared group had higher scores in both section 1 (22.1% versus 18.3%) and section 2 (13.7% versus 12.8%) compared to those in the Assessment-focused group. However, this difference was only statistically significant for section 1. This suggests that students who followed the intended structure of the flipped classroom performed better on assessments compared to those who prioritized assessments over class activities. Student activity was assessed by their frequency of accessing the class website and number of activities completed in each section. The differences in activity between the two groups were statistically significant for the entire semester, with the Prepared group having more sessions and total events than the Assessment-focused group. Although the Prepared group had a slightly lower number of events per session, it can be inferred that they accessed the virtual learning environment (VLE) for specific purposes each time.

When looking at activities specific to each section, it was observed that students in the Prepared group consistently accessed the VLE more frequently than those in the Assessment-focused group for both sections 1 and 2, with significantly higher numbers of sessions and events. Additionally, when examining average weekly sessions for each group over a 12-week semester, it was found that the Prepared group remained consistently more active throughout. Interestingly, even during week 8 when there was no class due to individual group meetings for the project, the Prepared group remained active on the VLE. On the other hand, activity levels for the Assessment-focused group declined during weeks 9 to 11 and increased again during the final week when assessments were approaching. This implies that this group engaged more with online materials as exams drew closer.

4.2 Improved Pattern Recognition Process

We present a method for categorizing student actions in the classroom through facial expressions, with the goal of determining their level of comprehension. Our approach involves training our model (Model α) on facial expressions from a large dataset (Dataset 1) and then fine-tuning it for use in classifying student behavior (Model β) using a smaller dataset (Dataset 2). We outline our proposed SLBPR-RNN approach, which consists of three main steps: (i) detecting and tracking faces, (ii) recognizing facial expressions, and (iii) classifying student behavior. Each step is explained further in the following sections. To detect faces in a video or image, we use the Haar Cascade algorithm developed by Viola and Jones, known for its accuracy in identifying one or more faces. In our implementation, we apply this algorithm to the first frame of a video using OpenCV, and then use the dlib library to track faces in subsequent frames. Once faces are detected and tracked, we extract them from each frame and resize them to 48 × 48 pixels while converting them to grayscale. Figure 5 provides an overview of our approach for classifying student behavior.

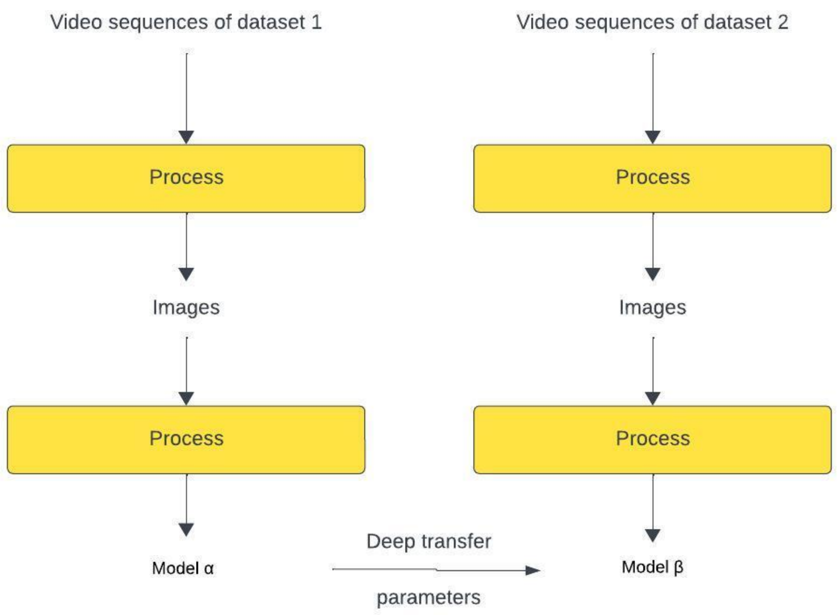

Figure 5 - An Overview of student's behavior classification approach.

For feature extraction, we adopt a per-trained model, VGG-16. Our team has made modifications to the VGG-16 architecture in order to improve its accuracy in classification tasks. This revised model includes 8 convolutional layers, 4 max-pooling layers, 1 flatten layer, 1 dropout layer, and 2 fully connected dense layers. To enhance the accuracy even further, we have incorporated the ELU activation function in all layers and utilized SAME padding in each convolutional layer to preserve the input size.

The specific structure of our modified VGG-16 model can be described as follows:

- The first and second convolutional layers (Conv1-1 and Conv1-2) consist of 32 feature kernel filters with a filter size of 3×3. This results in an output size of 48x48x32, which is then passed through a max-pooling layer with a stride of 2.
- The third and fourth convolutional layers (Conv2-1 and Conv2-2) utilize 64 feature kernel filters with a filter size of 3×3. A subsequent max-pooling layer with a stride of 2 reduces the output to 24x24x64.
- For the fifth and sixth convolutional layers (Conv3-1 and Conv3-2), we have employed 128 feature maps with a filter size of 3×3, resulting in an output size of 12x12x128. These are followed by another max-pooling layer with a stride of 2.
- The seventh and eighth convolutional layers (Conv4-1 and Conv4-2) consist of 256 kernel filters with a filter size of 3×3, followed by a max-pooling layer with a stride of 2.
- The flatten layer transforms the data into a one-dimensional array with a size of 2304.
- To prevent overfitting, we have incorporated a dropout layer with a rate of 0.2.
- The final layers consist of two fully connected hidden layers (Dense 1 and Dense2) each with 2304 units, followed by a soft max output layer.

In order to maximize the use of existing datasets and address the challenge of limited labeled data, we have also implemented deep transfer learning techniques. This involves utilizing pre-trained models, such as VGG-16 trained on Image Net, and transferring the learned weights to new tasks. In our case, we have used the pre-trained VGG-16 model as an efficient feature extractor for classifying student behavior in the classroom, even with a limited amount of video data from a separate dataset. We have transferred the learned weights from the facial expression recognition task on the first dataset to classify student behavior. The architecture of our modified VGG-16 model is illustrated in Figure 6.

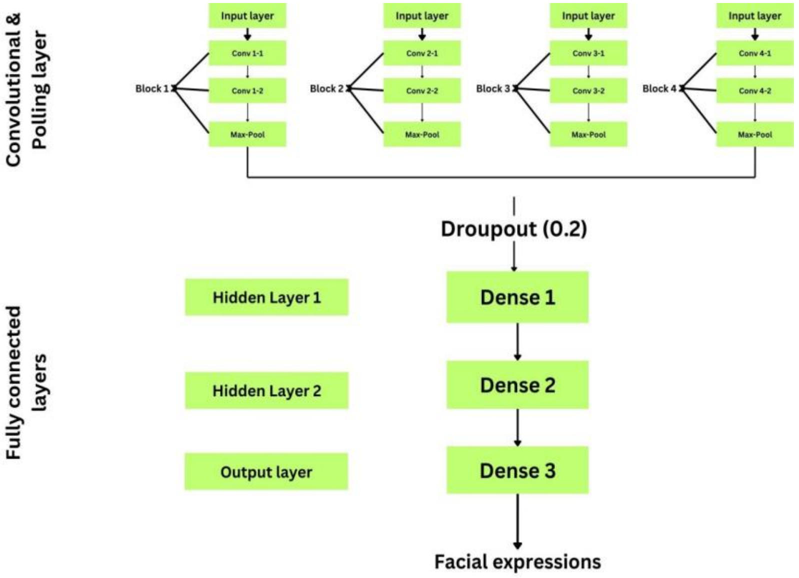

Figure 6 - Architecture of Modified VGG-16 Model.

Furthermore, we fine-tune the model for student behavior classification using data augmentation techniques to enhance accuracy and mitigate over fitting. This technique involves introducing small variations to the data-set, such as rotation, saturation, Gaussian blur, horizontal stretch, and desperation, without altering the central object. By augmenting the data-set, we increase the diversity of training examples, thus improving the robustness and generalization of the model.

V. PERFORMANCE ANALYSIS:

Simulation demonstration of the proposed SLBPR-RNN is performed in python with the presence of the SCB Student learning Behavior dataset [24]. The performance of the presented models is elaborated in this section in a detailed manner.

5.1 SCB Student learning Behavior dataset: These datasets are made up of a range of data gathered from students' engagements with educational platforms, systems, or materials. These collections of information may encompass User actions, Evaluation data, Demographic details and Usage trends. They are frequently utilized in educational studies, the creation of personalized learning systems, learning analytics, and other associated aims. In figure 7, some of the details about the SCB dataset are given which includes class differences, different learning stages and varying shooting angles.

Figure 7 - SCB-Dataset

5.2 SCB-Dataset annotation distribution: In datasets, annotations are generally used to indicate labels, tags, or extra details given to the data elements. The annotation distribution shows how these labels are spread out within the dataset. For instance, in a dataset of images labeled for identifying objects, the annotation distribution would reveal the number of instances labeled as "reading," "using phones," "hand raises," and "writing," as well as their distribution throughout the dataset. In figure 8, the SCB dataset annotation is described.

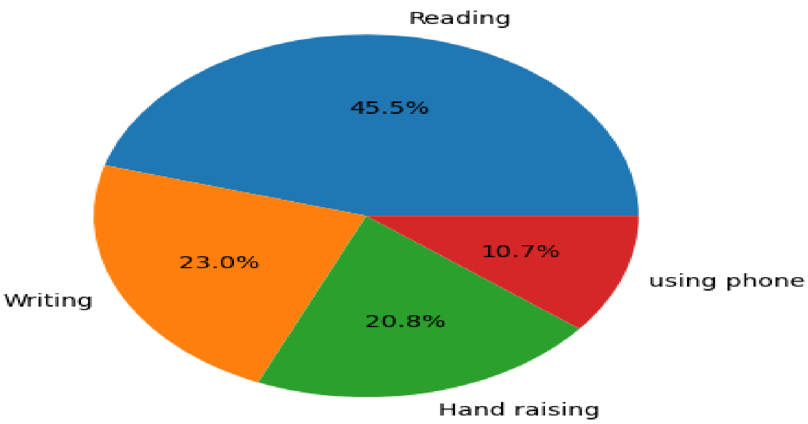

Figure 8 - SCB Dataset Annotation

5.3 Proposed SLBPR-RNN Training and Testing Accuracy: When using RNN for recognizing student learning behavior patterns in a flipped classroom, the terms "training accuracy" and "testing accuracy" refer to measures used for evaluating model's presentation during testing and training stages. Figure 9 illustrates the training and testing accuracy of the Proposed SLBPR-RNN model in the context of the SCB dataset.

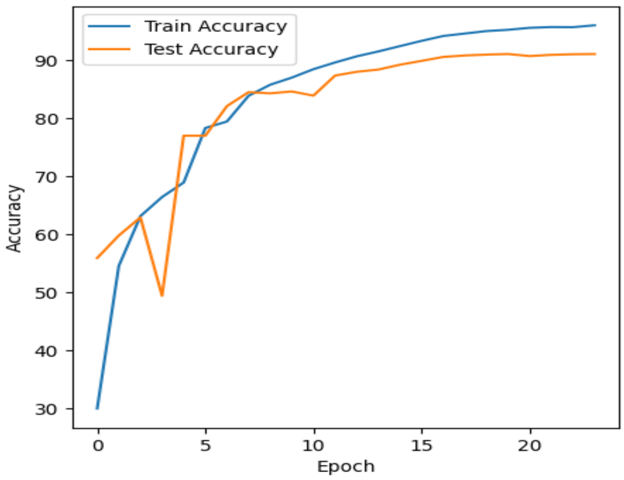

Figure 9 - SLBPR-RNN Training and Testing Accuracy

Training accuracy assesses how accurately the model predicts outcomes on the training data during the training stage. In the context of RNNs for recognizing student learning behavior patterns, training accuracy reflects how well the model adapts to our SCB dataset. It is determined by the percentage of correctly classified instances (or sequences) in the training SCB dataset. On the other hand, testing accuracy evaluates the model's predictions on new and unseen SCB data, also known as the testing SCB dataset. In RNNs for recognizing student learning behavior patterns, testing accuracy demonstrates how well the model can generalize to new and unseen SCB data. It is calculated in a similar manner to training accuracy but using a separate SCB dataset that was not used during training.

5.4 Comparative Analysis: The metrics which are used in the comparative analysis are Accuracy (%), Precision (%), Recall (%) and mAP@50%. The earlier baseline methods which are used for this comparative analysis are E-ELAN [25], YOLOv5 [26] and YOLOv7 [27] and as well it gets compared with the proposed SLBPR-RNN. Table 4 shows about the comparison of the presented methods with the considered metrics.

Table 4 – Comparative Performance

Methods	Class	Accuracy (%)	Precision (%)	Recall (%)	mAP@50%
E-ELAN	Reading	90.4	87.6	75.3	88.3
	Writing	81.1	84.2	83.3	89.1
	Hand raising	83.2	80.9	80.7	82.2
	Using phone	88.3	88.5	76.8	89.9
YOLOv5	Reading	93.6	87.6	75.3	78.3
	Writing	94.6	84.1	77.8	83.7
	Hand raising	85.8	79.4	86.9	92.6
	Using phone	88.1	81.2	72.4	89.4
YOLOv7	Reading	92.6	91.4	80.5	93.7
	Writing	90.1	92.9	83.2	94.9
	Hand raising	89.8	89	92.6	91.1
	Using phone	88.5	91.1	84.6	90.5
SLBPR-RNN	Reading	97.6	97.6	85.9	98.3
	Writing	94.1	94.1	82.1	99.1
	Hand raising	95.8	95.8	91.3	97.6
	Using phone	98.5	98.5	94.5	99.4

5.4.1 Accuracy Calculation: When using RNN for sequence modeling, the accuracy is determined by assessing the RNN's ability to accurately predict student learning behaviors in comparison to the correct labels. This is calculated by determining the percentage of correctly predicted behavior patterns out of all sequences in the dataset. It is typically represented as a percentage and can be mathematically calculated using formula (1).

$$Accuracy = \frac{No\ of\ Correctly\ Predicted\ Sequence}{Total\ No\ of\ Sequence} \times 100\% \tag{1}$$

The precision of RNNs in sequence modeling for detecting patterns in student learning behavior evaluates the model's accuracy in predicting behavior patterns, when compared to the actual labels. This metric offers a numerical indication of how well the model can identify and forecast significant patterns in student learning behaviors within a flipped classroom setting. In figure 10 illustrates the accuracy calculation of the methods like E-ELAN, YOLOv5, YOLOv7 and proposed SLBPR-RNN. The results show that the proposed SLBPR-RNN achieved maximum accuracy when compared with the earlier baseline methods.

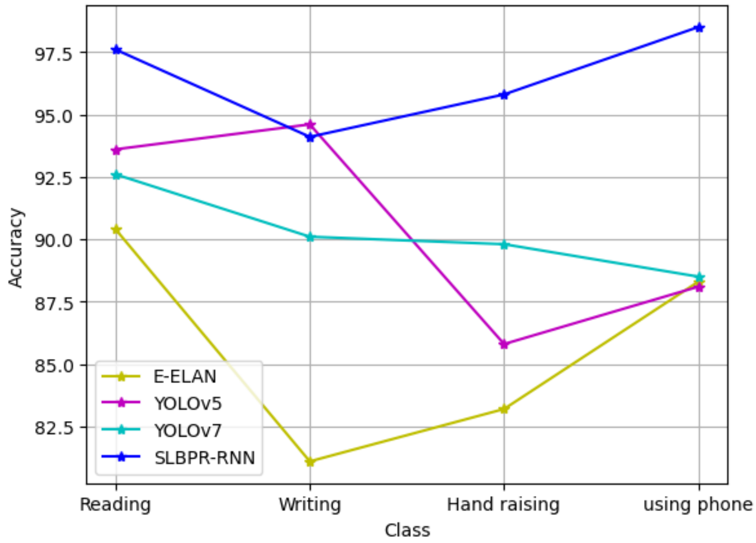

Figure 10 – Accuracy Calculation

5.4.2 Precision Calculation: When utilizing RNN for sequence modeling, the precision calculation is essential in determining the model's capacity to accurately predict positive outcomes. This measure is particularly valuable in evaluating the model's skill in recognizing distinct patterns of learning behavior among students through their interactions with educational materials. The formula (2) outlines the mathematical procedure for calculating precision.

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives + False\ Positives\ (FP)} \quad (2)$$

In the equation (2), the term TP refers to the sequences that the model accurately recognizes as being part of a specific behavior pattern, such as active engagement. On the other hand, FP refers to the sequences that are mistakenly identified by the model as being part of the behavior pattern when they are not. Precision is a measure that determines the accuracy of correctly identifying positive cases (true positives) in relation to all the cases the model identified as positive. It reflects the level of certainty we have that the instances classified as a specific behavior pattern by the model are correct. In recognizing student learning behavior patterns, precision assists in evaluating how effectively the model detects particular behaviors (e.g., active engagement) within a sequence of student interactions. A higher precision suggests that the model is making fewer incorrect predictions, which is advantageous when our goal is to accurately capture behavior patterns. In figure 11, the precision of the methods like E-ELAN, YOLOv5, YOLOv7 and proposed SLBPR-RNN are discussed and it proves the superiority of the proposed work.

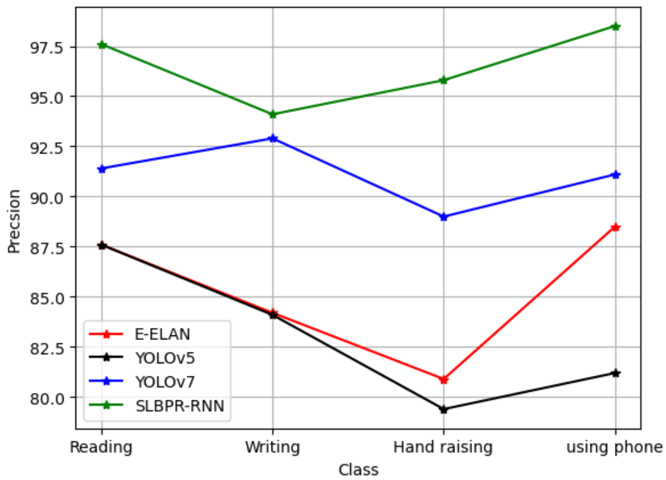

Figure 11 – Precision Calculation

5.4.3 Recall Calculation: When utilizing RNN for sequence modeling, it is crucial to calculate recall in order to evaluate the model's proficiency in accurately identifying all occurrences of a specific student learning behavior pattern within the student interactions with educational materials in a flipped learning setting. This calculation can be represented mathematically through equation (3).

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives + False\ Negatives\ (FN)} \quad (3)$$

In equation (3), the term True positive (TP) refers to the sequences that the model accurately recognizes as part of a particular behavior pattern, such as active engagement. False negative (FN) denotes the sequences that are mistakenly labeled by the model as not being part of the behavior pattern, although they truly are. Reminder is a metric that indicates the proportion of correctly identified positive cases among all actual positive cases. It illustrates the model's capability to detect all occurrences of a specific pattern of student learning behavior within a sequence of student interactions. In recognizing patterns of student learning behavior, reminder aids in comprehending how well the model recognizes and captures all occurrences of specific behaviors (e.g., active engagement) within student interactions. A higher reminder indicates that the model is effectively capturing a larger proportion of actual instances of the behavior pattern, which is important when we want to avoid missing any relevant instances. In figure 12 the performance of the Recall is shown for the methods such as E-ELAN, YOLOv5, YOLOv7 and proposed SLBPR-RNN were the effectiveness of the proposed work is proved.

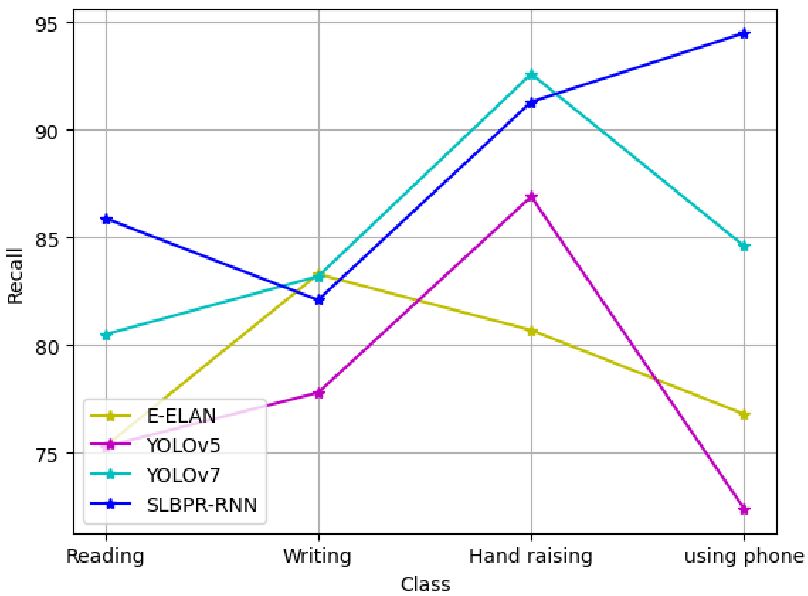

Figure 12 – Recall Calculation

5.4.4 mAP@50% Calculation: RNN sequence modeling utilizes the @50% calculation, also known as mean Average Precision at 50%, to assess how well the model can detect significant patterns in student learning behavior from their interactions with educational materials. The determination of Average Precision (AP) involves computing the area under the precision-recall curve (AP) for each behavior pattern. Mean Average Precision (mAP) is obtained by averaging the AP scores for all behavior patterns. When considering mAP@50%, only the precision values at a recall level of 50% are taken into account. This allows for evaluating precision at a specific recall level, in this case, 50%. The mAP@50% metric is valuable when there is a need for a fair balance between recall and precision. It serves as a calculate of effectiveness of the SLBPR-RNN model in detecting relevant patterns in student learning behavior, while also maintaining a 50% recall level. This metric is especially useful in applications where both precision and recall play crucial roles, such as studying student learning behavior in a FC environment. In figure 13, the performance of the presented methods like E-ELAN, YOLOv5, YOLOv7 and proposed SLBPR-RNN are calculated in terms of mAP@50% and the superiority of the proposed work is understood.

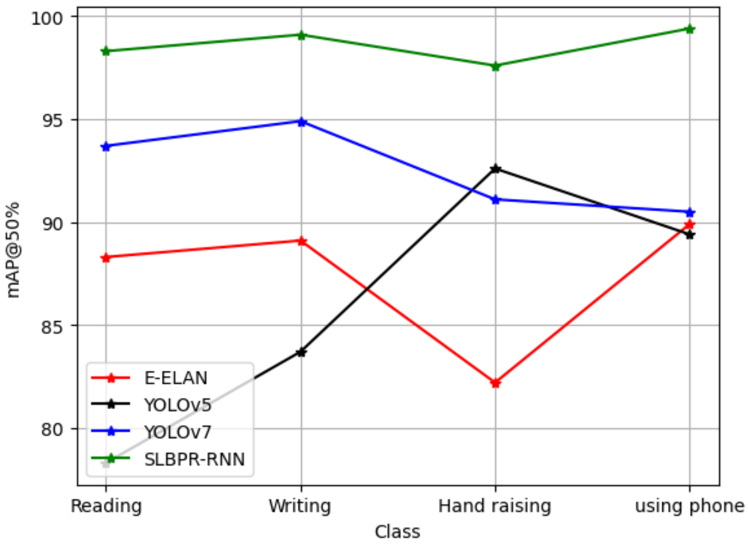

Figure 13 - mAP@50% Calculation

VI. CONCLUSION

Using RNN-based sequence modeling allows for the recognition and comprehension of unique patterns in learning behavior, opening up possibilities for personalized learning experiences. By tailoring educational materials to match individual preferences, pace, and comprehension styles, students can engage with content in a way that best suits them. The system also facilitates early detection of learning difficulties and gaps in understanding, providing educators with valuable insights to address potential issues before they worsen. This proactive approach supports a more effective and targeted intervention strategy. In a flipped classroom setting, where students access instructional material outside of regular class time, incorporating RNNs enhances the efficacy of both pre class activetie and in class activitie by adapting them to each student's learning patterns. This collaboration strengthens the flipped classroom model as a whole. The integration of RNNs in recognizing learning behavior presents opportunities for further research and innovation in education. This technology-driven approach highlights the potential of artificial intelligence to revolutionize the future of education.

REFERENCES

[1] Weijun Wang, “Application of deep learning algorithm in detecting and analyzing classroom behavior of art teaching”, Systems and Soft Computing, vol. 6, pp. 200082, 2024, doi: 10.1016/j.sasc.2024.200082

[2] Tom Adams, Bob Koster, et.al, “Patterns in student teachers' learning processes and outcomes of classroom management during their internship”, Teaching and Teacher Education, vol. 120, pp. 103891, 2022, doi: 10.1016/j.tate.2022.103891

[3] Adriano Bressane, Daniel Zwirn, et.al, “Understanding the role of study strategies and learning disabilities on student academic performance to enhance educational approaches: A proposal using artificial intelligence”, Computers and Education: Artificial Intelligence, vol. 6, pp. 100196, 2024, doi: 10.1016/j.caeai.2023.100196

[4] Beatriz Flamia Azevedo, Maria F. Pacheco, et.al, “Dataset of mathematics learning and assessment of higher education students using the MathE platform”, Data in Brief, vol. 53, pp. 110236, 2024, doi: 10.1016/j.dib.2024.110236

[5] Yunxiang Zheng , Junyi Zhang, et.al, “Effects of digital game-based learning on students’ digital etiquette literacy, learning motivations, and engagement”, Heliyon, vol. 10, pp. e23490, 2024, doi: 10.1016/j.heliyon.2023.e23490

[6] Taoufik Ben Abdallah, Islam Elleuch, et.al, “Student Behavior Recognition in Classroom using Deep Transfer Learning with VGG-16”, Procedia Computer Science, vol. 192, pp. 951–960, 2021, doi: 10.1016/j.procs.2021.08.098

[7] Miguel A.Conde, Francisco J. Rodríguez-Sedano, “Is learning analytics applicable and applied to education of students with intellectual/developmental disabilities? A systematic literature review”, Computers in Human Behavior, vol. 155, pp. 108184, 2024, doi: 10.1016/j.chb.2024.108184

[8] Joe Hazzam, Stephen Wilkins, “|The influences of lecturer charismatic leadership and technology use on student online engagement, learning performance, and satisfaction”, Computers & Education, vol. 200, pp. 104809, 2023, doi: 10.1016/j.compedu.2023.104809

[9] Swadha Gupta, Parteek Kumar, et.al, “EDFA: Ensemble deep CNN for assessing student’s cognitive state in adaptive online learning environments”, International Journal of Cognitive Computing in Engineering, vol. 4, pp. 373–387, 2023, doi: 10.1016/j.ijcce.2023.11.001

[10] Dagnew Gebrehiwot Giday, Elantheraiyan Perumal, "Students' perception of attending online learning sessions post-pandemic", Social Sciences & Humanities Open, vol. 9, pp. 100755, 2024, doi: 10.1016/j.ssaho.2023.100755

[11] Mariano Velamazan, Patricia Santos, et.al, "User anonymity versus identification in computer-supported collaborative learning: Comparing learners' preferences and behaviors",Computers & Education, vol. 203, pp. 104848, 2023, doi: 10.1016/j.compedu.2023.104848

[12] Mario A. Manzi-Puertas, Izaskun Agirre-Aramburu, et.al, "Navigating the student entrepreneurial journey: Dynamics and interplay of resourceful and innovative behavior", Journal of Business Research, vol. 174, pp. 114524, 2024, doi: 10.1016/j.jbusres.2024.114524

[13] Xinyi Dai, "A study on mindful agency's influence on college students' engagement with online teaching: The mediating roles of e-learning self-efficacy and self-regulation", Acta Psychologica, vol. 243, pp. 104146, 2024, doi: 10.1016/j.actpsy.2024.104146

[14] Navdeep Verma, Dr Seyum Getenet, et.al, "Designing an artificial intelligence tool to understand student engagement based on teacher's behaviours and movements in video conferencing", Computers and Education: Artificial Intelligence, vol. 5, pp. 100187, 2023, doi: 10.1016/j.caeai.2023.100187

[15] Musa Adekunle Ayanwale, Mdutshekelwa Ndlovu, "Investigating factors of students' behavioral intentions to adopt chatbot technologies in higher education: Perspective from expanded diffusion theory of innovation", Computers in Human Behavior Reports, vol. 14, pp. 100396, 2024, doi: 10.1016/j.chbr.2024.100396

[16] Maria Ripolles, Andreu Blesa, "The role of teaching methods and students' learning motivation in turning an environmental mindset into entrepreneurial actions", The International Journal of Management Education, vol. 22, pp. 100961, 20214, doi: 10.1016/j.ijme.2024.100961

[17] John N. Walsh, "Using cluster analysis to identify procrastination and student learning strategies in a flipped classroom", The International Journal of Management Education, vol. 22, pp. 100936, 2024, doi: 10.1016/j.ijme.2024.100936

[18] Kourosh Borhani, Richard T.K. Wong, "An artificial neural network for exploring the relationship between learning activities and students' performance", Decision Analytics Journal, vol. 9, pp. 100332, 2023, doi: 10.1016/j.dajour.2023.100332

[19] Xiaohui Tao, Aaron Shannon-Honson, "Towards an understanding of the engagement and emotional behaviour of MOOC students using sentiment and semantic features', Computers and Education: Artificial Intelligence, vol. 4, pp. 100116, 2023, doi: 10.1016/j.caeai.2022.100116

[20] Md Ochiuddin Miah, Umme Habiba, et.al, "ODL-BCI: Optimal deep learning model for brain-computer interface to classify students confusion via hyperparameter tuning", Brain Disorders, vol. 13, pp. 100121, 2024, doi: 10.1016/j.dscb.2024.100121

[21] I-Ching Chen, Gwo-Jen Hwang, et.al, "From design to reflection: Effects of peer-scoring and comments on students' behavioral patterns and learning outcomes in musical theater performance", Computers & Education, vol. 150, pp. 103856, 2020, doi: 10.1016/j.compedu.2020.103856

[22] Feng-Ying Li , Gwo-Jen Hwang, et.al, "Effects of a concept mapping-based two-tier test strategy on students' digital game-based learning performances and behavioral patterns", Computers & Education, vol. 173, pp. 104293, 2021, doi: 10.1016/j.compedu.2021.104293

[23] Surina He, Carrie Demmans Epp and Fu Chen, "Examining change in students' self-regulated learning patterns after a formative assessment using process mining techniques", Computers in Human Behavior, vol. 152, pp. 108061, 2024, doi: 10.1016/j.chb.2023.108061

[24] Yang, F and Wang, T, "SCB-Dataset3: A Benchmark for Detecting Student Classroom Behavior", arXiv preprint arXiv:2310.02522, 2023.

[25] Fan, Y, "SCB-dataset: a dataset for detecting student classroom behavior" arXiv preprint arXiv:2304.02488, 2023.

[26] Trabelsi, Z., Alnajjar, F, et.al, "Real-time attention monitoring system for classroom: A deep learning approach for student's behavior recognition", Big Data and Cognitive Computing, vol. 7, no. 1, pp. 48, 2023.

[27] Wang, Z., Li, et.al, "Student learning behavior recognition incorporating data augmentation with learning feature representation in smart classrooms", Sensors, vol. 23, no. 19, pp. 8190, 2023.