

Engagement Detection during E-Learning Classes using Machine Learning Algorithms

Prabhat Kashyap, Hemant Kumar, Nancy Jain

Department of cse, miet, meerut.

Abstract: Engagement detection during e-learning classes is super important because it helps us understand how students learn and stay interested in their studies. Lately, computers have been getting really good at figuring out if students are paying attention during online classes. In this research, we're using facial expressions and eye movements to see if students are focused during online classes. We're comparing different computer programs to figure out which ones are best at doing this job. We're also looking at how different features, like facial expressions and where students are looking, affect how well these programs work.

Student engagement refers to students' psychological involvement in the acquiring and comprehending academic work tries to teach you stuff, like what you know, what you can do, and the skills you have. In realm of education, student engagement stands as a important factor influencing academic success and overall learning outcomes. Student accomplishment is directly proportionate to his or her level of engagement. The main issue we face here is that it is not feasible for the teacher to teach in the online classroom and simultaneously notice which students are distracted and which are not. Distracted students not only fail to learn themselves but are also much more likely to cause disturbances during the class. . By using machine learning algorithms to detect engagement levels during e-learning classes, educators can identify students who may be struggling and provide them with additional support. Our study also highlights the importance of certain features for engagement detection during e-learning classes, which can inform the design of more effective e-learning environments.

Our study helps us understand how computer programs can tell if students are paying attention in online classes. This helps make online learning better in the future.

Model was developed to identify and inform inattentive learners. We must get the input data from this video stream before doing the analysis, and the model must be as successful under dim lighting

Faculty -Jagbeer Singh, Department of CSE, MIET, Meerut.

[1] Create a mechanism to measure students levels of involvement based on their eye movements.

[2] Based on the amount of participation classify into three groups: "extremely involved," "usually engaged," and "not engaged"

Key Words : Machine Learning, CNN Module, Viola & Jones algorithm, HaarCas- cade Algorithm, media pipe , Arriaga, Valdenegro-Toro, and Ploger , facial emotion classification, mini-Xception .

Introduction

1.1 Mentioning the introduction of project.

Studying how much students are paying attention in online classes is really important because it helps us learn more about how students learn and stay interested. Lately, computers have been used more and more to figure out if students are engaged during these online classes.

This study aims to see how well different computer programs can tell if students are paying attention during online classes. We're going to compare how these programs work with different settings. We'll also check out how certain things, like facial expressions and where students look, affect how well these programs work.

The rest of this paper is set up like this: First, in Section 1.2, we talk about what other researchers have found out about how to tell if students are paying attention in online classes and how computers can help with that. Then, in Section III, we explain how we did our study, like where we got our data from, how we got it ready to use, and what computer programs we used. In Section IV, we share what we found out from our study, like which computer programs worked best and how different things, like facial expressions, affected the results. After that, in Section V, we talk about what our findings mean for teachers and other researchers who want to make online classes better for students. Finally, in Section VI, we sum up the most important things we learned and give some ideas for what future research could look into.

1.2 Background history related to project.

Since the intervening-1990s, the idea of 'Student Engagement' has accumulate significant consideration, seeking allure ancestries back to Astin's 1984 paper on junior engrossment. Following in the footsteps of 'The Student Experience' and 'Research-surpassed Teaching,' 'Student Engagement' has arose as a focus for reconstructing university everywhere. The literature usually links pupil engrossment in intense instructional activities to helpful consequences to a degree vindication, steadfastness, academic realization, and social date. Given the current business-related challenges confronted by Higher Education Institutions, inviting, maintaining, and cultivating students is critical. Kuh (2003) stresses that juniors' accomplishment is more affected by their activities all along their opportunity as graduates than their culture or place they study. If 'Student Engagement' lives until its promises, maybe the key to accomplishing these aims.

Alkabbany, I., Ali, A., Farag, A., Bennett, I., Ghanoum, M., Farag, A.: Measuring undergraduate date level utilizing facial news. In: 2019 IEEE International Conference on Im- age Processing (ICIP. pp. 3337–3341. IEEE (2019)

.Engagement goes further absolute difficulty or partnership; it encompasses sentiments, understanding, and project (Harper & Quayle, 2009 (a), 5). Acting outside a sense of date is similar to agreement, while feeling busy outside operation leads to detachment. Fredricks, Blumenfeld & Paris (2004, 62-3) climax three ranges of student date, suitable from Bloom (1956):

1. Behavioural Engagement: Students the one are behaviourally operating and adhere to averages, presenting attendance, difficulty, and a lack of causing trouble behaviour.2. Emotional Engagement: Emotionally engaged juniors happening concerning feelings and intuition responses such as interest, possession, or sense of relationship.3. Cognitive Engagement: Cognitively operating pupils invest in their education, surpass necessities, and relish challenges.

Literature Review:

Dewan et al. [1] presents an overview of the machine learning in involvement detection in online learning. The model's framework is made up of five main modules: face detection from camera, landmark extraction, tracking, classification in various involvement level, and conclusion.

Manseras and others. [2] presents a method for categorizing students' difficulty first characteristics Tests and compares various classification techniques so construct the optimal model. SVM accuracy outperforms the different two classification algorithms as a whole of three datasets. Mohamad Nezami and others. [3] shows a deep learning model that increases partnership identification from photographs by pre-preparation on effortlessly available fundamental first verbalization dossier before training on specialized connection lineaments

1.3 Supported technologies, algorithms helped in project development.

The successful development of the project was made possible through the adept utilization of cutting-edge technologies and sophisticated algorithms. Machine learning played a pivotal role, employing the Convolutional Neural Network (CNN) module for intricate tasks. The Viola & Jones algorithm and HaarCascade Algorithm were integral components in the detection system, alongside MediaPipe and the approach by Arriaga, Valdenegro-Toro, and Ploger for facial emotion classification. The mini-Xception model, incorporating pooling to reduce size, was employed for nuanced emotion categorization into seven distinct states: Angry, Disgust, Fear, Happy, Sad, Surprise, or Neutral. The aggregation index, a critical rhythmical, was computed based on the assurance worth and the particular aggregation index of the dominant empathy. Two important algorithms, the HaarCascade Algorithm and a Convolution Neural Network, synergistically provided to the adeptness of the discovery system. The concluding treasure surpassed in detecting the graduate's facade face in the image and just location analysis domain within the face. Additionally, a wisdom-intelligent and breakable Conv2D tier was influential in the depth guess process. The aggregation index (CI) forecast was carefully brought into harmony based on moving shadings, handling a rule that dynamically adjusted the CI principles for various fervors, guaranteeing an inclusive assessment of undergraduate date.

Proposed Work Plan:

2.1 General architecture/ Flow chart/ DFD of overall system to be designed.

The envisioned system adopts a sophisticated architecture, seamlessly integrating various modules to accomplish the overarching goal of efficient attention detection and emotional classification. The system's architecture is illustrated through a comprehensive flow chart and Data Flow Diagram (DFD), elucidating the intricate interactions between different components.

System Overview & Methodology

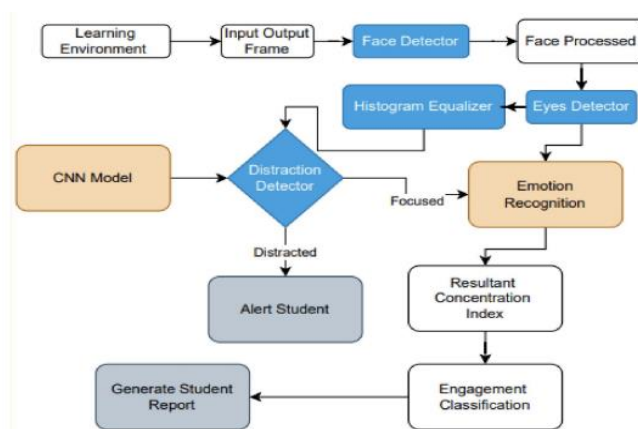


Figure 1: System Flowchart

2.2 Description of various modules of the system:

The system comprises several meticulously designed modules, each serving a specific function in the holistic approach to attention detection. These modules include the image preprocessing module, emotion classification module, and the core attention detection module. The image preprocessing module ensures that the input data is appropriately formatted and standardized for subsequent stages. The emotion classification module employs advanced algorithms, such as the mini-Xception model, to accurately categorize facial expressions into distinct emotional states. . The attention detection module, combining the HaarCascade Algorithm and a Convolution Neural Network, energetically identifies and evaluates junior data based on first countenance and touching cues.

The indicator engagement scheme holds three modules: -

[1] Distraction Detector: Using the student's face categorize the welcome/her consideration state in two together types "Distracted" or "Focused". Leveraging first features, this module provides real-time insights into the student's engagement level, crucial for understanding the learning environment.

[2] Facial Emotion Recognition: Classify the first verbalization in an individual of the seven classifications: Angry, Disgusted, Fearful, Happy, Sad, startled, or Neutral. By employing advanced algorithms, the system accurately captures and interprets the subtle nuances of a student's emotional states, contributing to a nuanced understanding of their engagement.

[3] Engagement Classification: Calculate the aggregation index and categorize in individual of three types: Very Engaged, Nominally Engaged and Not Engaged.

Listed below are the primary emotions and their corresponding intensities:

Table 1

Dominant Emotion	Weightage
Normal	0.8
Happy	0.7
Surprised	0.5
Sad	0.4
distaste	0.3
Anger	0.27
Fearful	0.28

[1] Very Involved: when value of CI is more than 0.5 then that student lies in Very Engaged category.

[2] Normally Involved: when value of CI is less than equal to 0.5 then the student lies is Normally Engaged category.

[3] Distracted: when first CNN model given 0 output then that student is Distracted.

2.3 Algorithm of main complement of the system:

The gist of bureaucracy depends a carefully created algorithm that seamlessly integrates the HaarCascade Algorithm and a Convolution Neural Network. This treasure surpasses in detecting and resolving the junior's first verbalizations and data levels. Additionally, a depth-wise and separable Conv2D layer plays a crucial role in depth estimation, contributing to the system's overall accuracy. The concentration index (CI) computation algorithm, intricately adjusted based on emotional nuances, ensures a nuanced evaluation of student engagement, providing valuable insights into the learning experience.

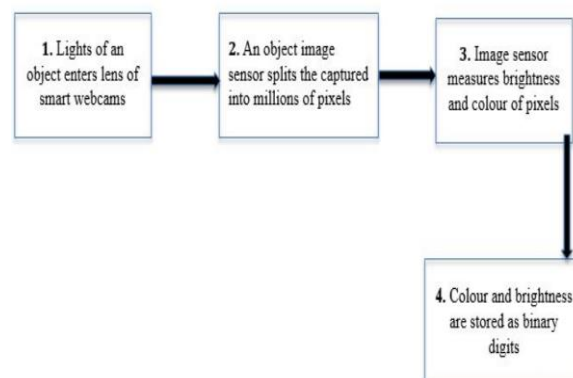


Figure 2: Face Detector

- [1] Open-face Model for Facial Recognition: For Face Recognition Open face library is used which uses a pre-trained model of Dlib and an open CV. After Face Recognition facial landmarks are extracted and from all facial landmarks choose 8 landmarks (68,107,236,117,336,298,346,456) which identify the eye region which is the region of interest here.
- [2] Histogram Equalization is like tweaking a photo to make it look better. It's a way to adjust the brightness of different parts of the picture so it stands out more. This often makes photos look clearer and sharper, especially when they have similar brightness level.

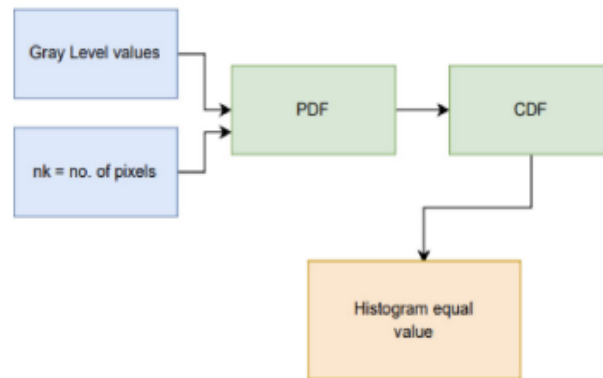


Figure 3: Workflow of Histogram Equalizer

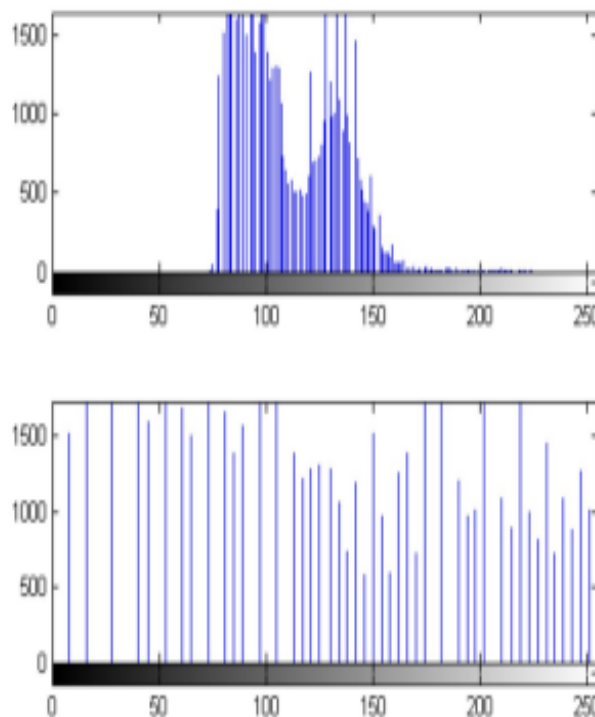
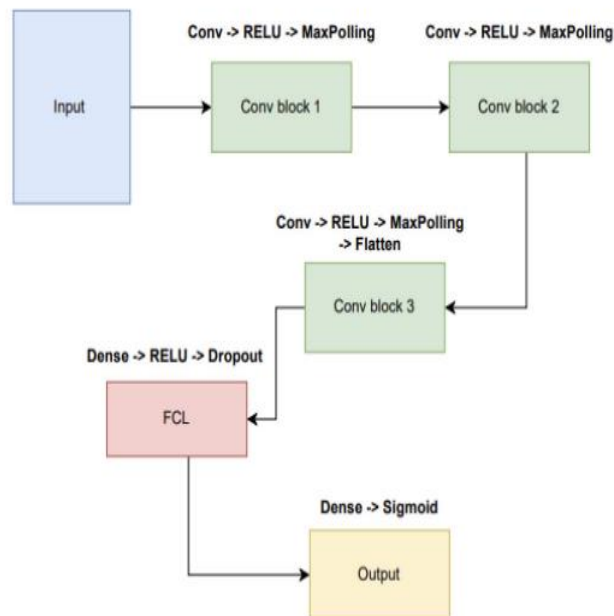


Figure 4: Histogram Equalization Comparison Graph

CNN Model for Binary Classification:**Figure 5 : Work-flow diagram CNN Model**

- [1] CNN model for Emotion Emotion: Second CNN architecture for facial emotion detection as stated above using Python that can recognize emotions on people's faces. We used different tools like NumPy, Keras, and OpenCV to help us do this.
- [2] The Concentration Index (CI) is figured out by multiplying the Dominant Emotion Probability (DEP) value by its matching Emotion Weight (EW). It's like getting a score by multiplying how likely a certain emotion is by how much it matters. Example, if emotion of person is 'neutral,' the CI is calculated as $(DEP \times 0.9) \times 100$. In contrast, for 'happy,' the formula is $(DEP \times 0.6) \times 100$, and so forth.
- [1] If the emotion is 'neutral', then the Concentration Index (CI) is calculated by multiplying the Dominant Emotion Probability (DEP) by 0.9 and then by 100.
- [2] If the emotion is 'happy', then the CI is calculated by multiplying the DEP by 0.6 and then by 100.
- [3] If the emotion is 'surprised', then the CI is calculated by multiplying the DEP by 0.5 and then by 100.
- [4] If the emotion is 'sad', then the CI is calculated by multiplying the DEP by 0.3 and then by 100.
- [5] If the emotion is 'scared', then the CI is calculated by multiplying the DEP by 0.3 and then by 100.
- [6] If the emotion is 'angry', then the CI is calculated by multiplying the DEP by 0.25 and then by 100.
- [7] If the emotion is 'disgust', then the CI is calculated by multiplying the DEP by 0.2 and then by 100. else CI = 0 (i.e., Distracted), This approach allows for a detailed and dynamic evaluation of concentration levels.

The algorithmic framework is comprehensive, encompassing the detection of facial features, emotion classification, and concentration index calculation. It integrates advanced techniques such as depth-wise and separable Conv2D layers and considers specific formulas for varied emotional states, ensuring a robust and accurate analysis of student engagement. The choice of activation functions, such as sigmoid and softmax, aligns with the specific requirements of binary and multi-class classification problems, contributing to the overall efficacy of the system.

3. Experimental Result Analysis:

3.1 Description of data set used.

In the pursuit of training the CNN for our engagement-detection system, a meticulously curated dataset from a Kaggle challenge was employed. This dataset is comprised of grayscale images of faces, each measuring 48×48 pixels, and is abundant with a diverse array of examples, totaling 35,887 instances. The richness of this dataset is instrumental in ensuring the robustness and generalization of the CNN, allowing it to effectively learn intricate patterns and nuances associated with facial expressions and engagement states. Leveraging this Kaggle dataset provides a solid foundation for the training process, enhancing the system's ability to discern subtle variations in emotions and attention levels across a spectrum of real-world scenarios.

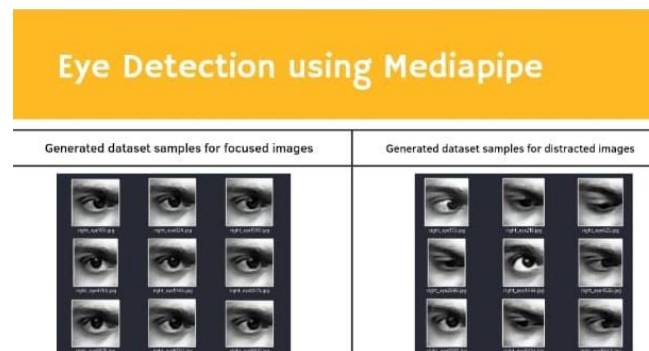


Figure 6: Test Data Details

- The dataset includes pictures of eyes that are 48 pixels wide and 48 pixels tall, and they're all in shades of gray.
- Two videos from each person were recorded for distraction and focused level respectively and from each video 25 random frames were taken.
- The data includes pictures of faces that are 48 pixels wide and 48 pixels tall, and they're all in shades of gray. The faces are positioned so that they're roughly in the middle of each picture and take up about the same amount of space.
- The goal is to figure out what emotion each face is showing. There are seven possible categories: Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral.

3.2 Evaluation of System Efficiency and Accuracy

The efficiency and accuracy of the designed engagement detection system were systematically evaluated based on specific parameters crucial for performance assessment. Two key factors, namely the effect of dropout on accuracy using the activation function of ReLU and the effect of dropout on accuracy using the hyperbolic tangent (tanh) activation function, were thoroughly investigated. Dropout, a regularization technique, was systematically applied to discern its impact on the system's precision under different activation functions.

Effect of dropout on accuracy using Relu Activation Function

```

21
22 model = Sequential()
23 model.add(Conv2D(32, (3, 3), input_shape=input_shape))
24 model.add(Activation('relu'))
25 model.add(MaxPooling2D(pool_size=(2, 2)))
26
27 model.add(Conv2D(32, (3, 3)))
28 model.add(Activation('relu'))
29 model.add(MaxPooling2D(pool_size=(2, 2)))
30
31 model.add(Conv2D(64, (3, 3)))
32 model.add(Activation('relu'))
33 model.add(MaxPooling2D(pool_size=(2, 2)))
34
35 model.add(Flatten())
36 model.add(Dense(64))
37 model.add(Activation('relu'))
38 model.add(Dropout(0.3))
39 model.add(Dense(1))
40 model.add(Activation('sigmoid'))
41
42 model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])

```

Figure 7

Additionally, the system's robustness was assessed by varying the filter size while employing the ReLU activation function. This exploration aimed to understand how changes in the filter size influence the accuracy of the system. By systematically varying this parameter, insights were gained into the optimal filter size for achieving the highest accuracy in facial emotion recognition and engagement classification. These comprehensive evaluations provide a nuanced understanding of the system's performance under various configurations, offering valuable insights for further optimization and refinement.

VARYING FILTER SIZE ON RELU ACTIVATION FUNCTION

```

model = Sequential()
model.add(Conv2D(32, (5, 5), input_shape=input_shape))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))

model.add(Conv2D(32, (5, 5)))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))

model.add(Conv2D(64, (5, 5)))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))

model.add(Flatten())
model.add(Dense(64))
model.add(Activation('relu'))
model.add(Dropout(0.3))
model.add(Dense(1))
model.add(Activation('sigmoid'))

model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])

```

Figure 8

Conclusion

In conclusion, our project has demonstrated the effectiveness of a Convolutional Neural Network (CNN) architecture for the binary classification of student engagement, successfully discerning nuanced states of attention as "Focused" or "Distracted."

Through meticulous experimentation, we conducted a comparative analysis of two prominent activation functions, namely ReLU and Tanh, in conjunction with varying dropout rates. The results underscored the significance of

hyperparameter fine-tuning, with ReLU coupled with a dropout rate of 0.3 emerging as the optimal configuration for achieving superior performance.

The investigation into different filter sizes further enriched our understanding of the system's intricacies. Notably, our findings indicated that 3x3 filters outperformed their 5x5 counterparts, providing more precise outcomes in the context of facial emotion recognition and engagement classification.

[1] This comprehensive exploration has not only advanced our comprehension of neural network architecture intricacies but has also significantly contributed to the refinement and efficacy of our engagement detection system. The fine-tuned parameters and insights gained from this analysis lay a solid foundation for the continued enhancement and real-world applicability of our innovative approach to student engagement assessment.

[1.1] "Real-Time Student-Engagement Detection during E-learning using emotion analysis and eye tracking" represents an innovative and timely initiative to transform the way we assess and enhance online education. With the potential to improve both the academic experience and the competitiveness of e-learning platforms, this project holds great promise in shaping the future of digital education. We anticipate that our research will not only benefit educators and learners but also contribute valuable knowledge to the broader fields of educational technology and human-computer interaction.

Scope of the Project

[1] Getting instant feedback from the system will allow the teacher to adjust the learning materials to better suit the student's needs.

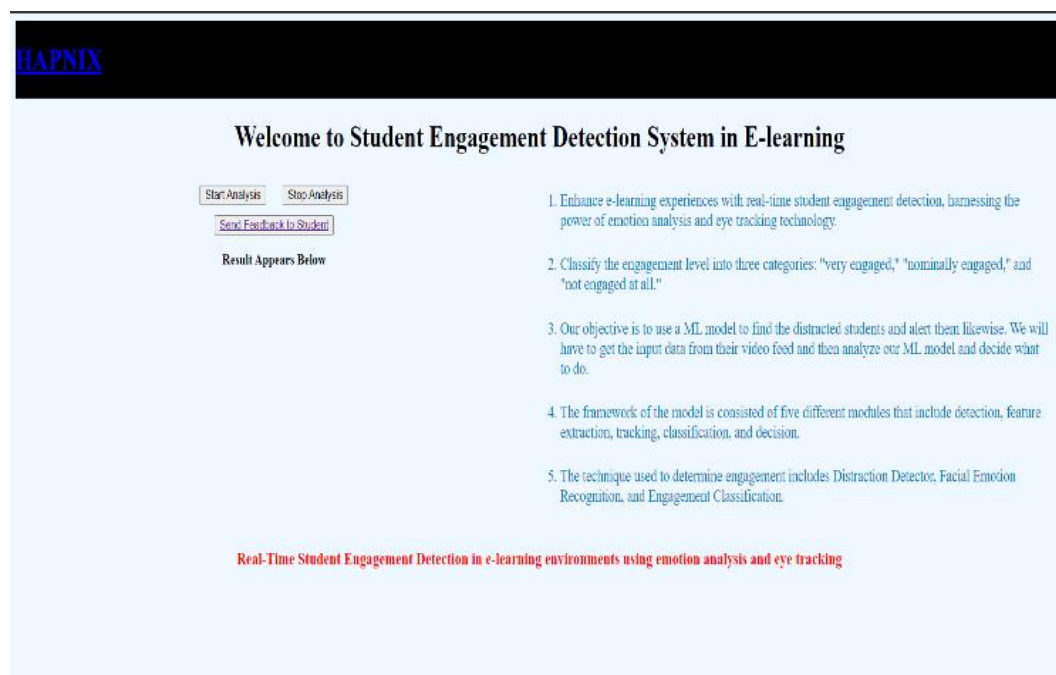


Figure 9

[2] At the end of each session, a report of each student concentration level will be generated.

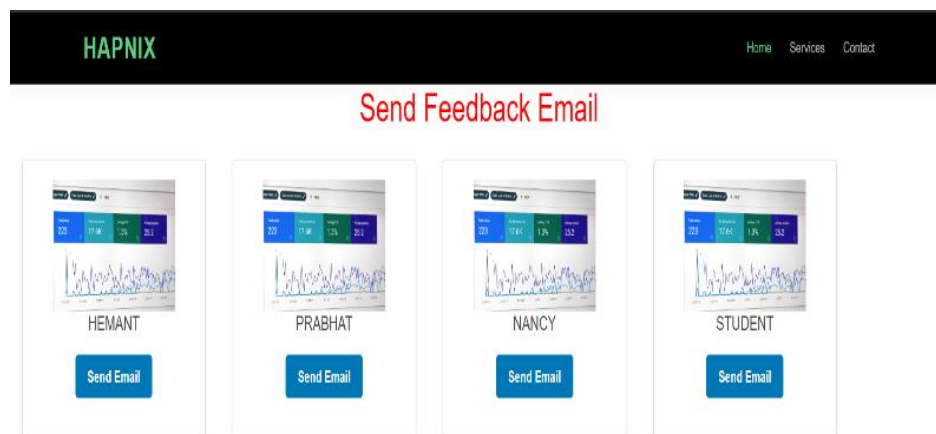


Figure 10

References:

- [1] Dewan, M., Murshed, M. and Lin, F.: 2019, Engagement detection in online learning: a review, *Smart Learning Environments* 6(1), 1–20.
- [2] Manseras, R., Palaoag, T. and Malicdem, A.: 2017, Class engagement analyzer using facial feature classification, no. November pp. 1052–1056.
- [3] Mohamad Nezami, O., Dras, M., Hamey, L., Richards, D., Wan, S. and Paris, C.: 2020, Automatic recognition of student engagement using deep learning and facial expression, *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer, pp. 273–289
- [4] Bosch, N. Detecting student engagement: Human versus machine. In *Proceedings of the 24th International Conference on User Modelling, Adaptation and Personalization*, Halifax, NS, Canada, 13–17 July 2016; pp. 317–320
- [5] Murshed, M.; Dewan, M.A.A.; Lin, F.; Wen, D. Engagement detection in e-learning environments using convolutional neural networks. In *Proceedings of the 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech)*, Fukuoka, Japan, 5–8 August 2019; pp. 80–86.
- [6] Wells, M.; Wollenschlaeger, A.; Lefevre, D.; Magoulas, G.D.; Poulouvassilis, A. Analysing engagement in an online management programme and implications for course design. In *Proceedings of the Sixth International Conference on Learning Analytics Knowledge*, New York, NY, USA, 25–29 April 2016; pp. 236–240.
- [7] Ashwin, S.T.; Guddeti, R.M.R. Automatic detection of students’ affective states in classroom environment using hybrid convolutional neural networks. *Educ. Inf. Technical*
- [8] H. Monkaresi, N. Bosch, R. A. Calvo, and S. K. D’Mello, “Automated detection of engagement using video-based estimation of facial expressions and heart rate,” *IEEE Transactions on Affective Computing*, vol. 8, no. 1, pp. 15–28, 2016.
- [9] J. Newnham, *Machine Learning with Core ML: An iOS developer’s guide to implementing machine learning in mobile apps*. Packt Publishing Ltd, 2018.

- [10] O. M. Nezami, M. Dras, L. Hamey, D. Richards, S. Wan, and C. Paris, "Automatic recognition of student engagement using deep learning and facial expression," in Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 2019, pp. 273–289.
- [11] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," arXiv preprint arXiv:1811.03378, 2018.
- [12] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," arXiv preprint arXiv:1511.08458, 2015.
- [13] S.-C. Wang, "Artificial neural network," in Interdisciplinary computing in java programming. Springer, 2003, pp. 81–100.
- [14] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," Journal of Big data, vol. 3, no. 1, pp. 1–40, 2016.
- [15] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan, "The faces of engagement: Automatic recognition of student engagement from facial expressions," IEEE Transactions on Affective Computing, vol. 5, no. 1, pp. 86–98, 2014.
- [16] M. A. Wiering and M. Van Otterlo, "Reinforcement learning," Adaptation, learning, and optimization, vol. 12, no. 3, 2012.
- [17] Y.-c. Wu and J.-w. Feng, "Development and application of artificial neural network," Wireless Personal Communications, vol. 102, no. 2, pp. 1645–1656, 2018.
- [18] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Metaface learning for sparse representation based face recognition," in 2010 IEEE International Conference on Image Processing. IEEE, 2010, pp. 1601–1604.