NOLEGEIN Journal of
Business Ethics, Ethos and CSR

ISSN: 2581-5601
Volume 5, Issue 2, 2022
DOI (Journal): 10.37591/NJBEEC

http://mbajournals.in/index.php/JoBEC/index

Review NJBEEC

Recognizing Student's Emotions in a Lecture

Zaheen B. Maruf*

Abstract

The realm of automatic emotion recognition based on facial expressions is a fascinating field of study that has been applied in numerous industries, including security, health care, and human–machine interface. In order to enhance computer predictions, researchers are exploring ways to decode, analyze, and extract facial expression features. With the remarkable achievements of Deep Learning, different types of architectures are being employed to boost performance. This article aims to present the latest research on Facial Expression Recognition using deep learning for the automatic identification of emotions. The article's objective is to assist researchers by examining current research and offering insights on how to enhance this field.

Keywords: Emotion recognition, facial expressions, interaction, learning technologies, psychology

INTRODUCTION

Teaching is an art that is unique in its own field. The ability of a certain student grasping a particular topic varies from student to student. Likely, as psychology says, that if a human is interested in the speaker, certain emotions will be reflected out on their face, but again these emotions are just vaguely different from person to person. In 2020-2021 Pandemic, as the physical classes shifted to a new method, which was certainly online, and that affected the student's attention span. Forecasting a student's attention skill in any certain lecture cannot be judged on a list of certain predefined parameters.

Certainly, Psychology presents some of the traits that a human show when someone is interested in listening out to the speaker:

- Asking Lots of Questions
- Body Contact
- Eye Contact

Contrary to the above, when a student is particularly not interested, he/she/they might reflect some of the emotions mentioned below:

- Yawning
- Sleeping
- Talking
- Using Mobile Phone

***Author for Correspondence**
Zaheen B. Maruf
E-mail: zaheenmaruf08@gmail.com

Research Scholar, Thakur Institute of Management Studies, Career Development and Research (TIMSCDR) Mumbai, India

Received Date: March 17, 2023
Accepted Date: March 26, 2023
Published Date: April 12, 2023

Citation: Zaheen B. Maruf. Recognizing Student's Emotions in a Lecture. NOLEGEIN Journal of Business Ethics, Ethos and CSR. 2022; 5(2): 28–38p.

All efforts of the instructor and his or her teaching aids became futile. To overcome these problems, many interactive learning technologies have been introduced in recent decades but have not yet improved but these can find a better way to analyze emotions in real time in an online learning environment. So, it is necessary for a system that analyzes student's moods in an online learning environment.

The consequences of this metric can also provide clues for instructors to make the topic interesting to

students. need. There is a question that can be raised in mind that this can be achieved using written comments. system, but this response will not be real-time.

Therefore, the goal of the system is to improve efficiency in an online learning/physical environment with the help of student emotions and also get real-time feedback in the form of feedback.

By assigning up an emotion recognition system, the system will be responsible for detecting and recognizing the student's emption, if the emotions fall toward the Sleeping, Yawning and Talking emotion spectrum, that clearly indicates that the teacher/professor is unable to grab student's attention.

To Look on the other side, if the student has a straight face and is speaking with the teacher, that certainly impacts that the teacher's lecture is interesting.

RELATED SYSTEM

Face Recognition

Face recognition [1] is a fascinating area of research that spans across multiple fields and disciplines. It is considered to be a fundamental aspect of human behavior, allowing for effective communication and interaction among people. In addition to its practical applications in areas such as bankcard identification, access control, Mugshots searching, security monitoring, and surveillance systems, the ability to recognize faces has become increasingly important in our daily lives. For instance, Facebook's algorithms have demonstrated an impressive ability to recognize faces with 98% precision, even after only a few instances of tagging, which is almost as accurate as humans.

Face Landmark Estimation

The concept involves identifying 68 specific points [2], commonly referred to as "landmarks," that are present on every human face. These landmarks include, but are not limited to, the outer edges of each eye, the inner edges of each eyebrow, and the top of the chin.

Facial Emotion Recognition

Facial Emotion Recognition [3], also known as FER, is a highly researched area within the field of computer vision that continues to undergo constant evaluation and improvement. The key question behind this area of study is whether a machine can accurately detect a person's emotions through their facial expressions. While humans are naturally adept at this task and perform it thousands of times per day, the challenge lies in training machines to achieve human-level performance in recognizing emotions using data sets. Recent advancements in computer vision have led to significant progress in this area, making these tasks much less challenging and requiring less effort to achieve good performance.

PROBLEM DEFINITION

The proposed system aims to provide continuous monitoring of students' concentration levels through the use of head rotation and eye movement tracking. By detecting and analyzing specific facial features, such as the direction of the student's gaze, the system can determine if the student is focused on the visual content (i.e., the learning material). The system then measures the student's head movements over time and uses this information to assess their concentration level.

To accomplish this, the system first recognizes the student's eyes, which serve as an important indicator of their level of interest and concentration. However, in cases where students become disengaged and look away from the camera, the system must use other features to accurately assess their concentration level. The outputs of these two components are combined and analyzed to provide a more comprehensive understanding of the student's level of engagement.

The proposed system architecture comprises several stages. The first two stages involve capturing and processing video input, which is then fed into the densitometry module. This module receives input

from the various submodules and generates feedback in the form of recorded graphs, providing insights into the student's concentration level and the duration of their focus.

System Design

The proposed system has two modules, an eye detection module and a head rotation module. The algorithms used in the proposed system are based on the Viola–Jones algorithm and LBP.

Eye Recognition Module

To detect the eye condition, the eye condition can be identified by first detecting the face from the captured frame and then detecting the eye area.

OBJECTIVE/SCOPE

- Identify the reflected emotions.
- Store the emotions extracted into a file for analysis.
- Analyze certain subjects as interesting or boring based on emotions gathered.
- Predict if students were attentive in a certain lecture.
- Scrutinize the teaching methods.
- Try and test different teaching methods.

ALGORITHMS AND MODELS.

In any machine learning or deep learning project, the data used to train the model is critical to its success. The quality of the data is directly reflected in the performance of the trained model. If the data is representative of the real world, then the model is more likely to perform well in real-world situations [4]. On the other hand, if the data contains many outliers, the model may produce irregular values in production. Therefore, it is crucial to ensure that the data used in machine learning or deep learning projects is of the highest quality, as it serves as the foundation upon which the rest of the project is built.

The Input Dataset

CK+48 data set contains directories of images and the directories are named after the emotion they contain. Within each directory there are a bunch of images for that particular emotion class.

OpenCV

OpenCV, or Open Source Computer Vision Library, is a software library for computer vision and machine learning that is available as a free and open-source tool. It was developed with the intention of providing a common infrastructure for computer vision applications and facilitating the incorporation of machine perception into commercial products. The use of the Apache 2 license makes it simple for businesses to utilize and modify the code as needed [5].

OpenCV offers over 2500 streamlined algorithms that encompass a broad range of classic and cutting-edge computer vision and machine learning techniques. These algorithms enable users to perform a variety of tasks, such as detecting faces, identifying objects, classifying human actions in videos, tracking camera movements, tracking moving objects, extracting 3D models of objects, generating 3D point clouds from stereo cameras, stitching images together to create high-resolution images of entire scenes, finding similar images in an image database, removing red eyes from flash images, following eye movements, recognizing scenery, and establishing markers to overlay.

The OpenCV user community is extensive, with over 47 thousand members, and the library has been downloaded over 18 million times [6].

Haarcascade

Figure 1 and 2 Haarcascade frontal face alt is an XML file that contains a serialized Haarcascade detector of faces (Viola–Jones algorithm) in the OpenCV library. It is a coded list of decision trees in

which each vertex tests one Haarcascade Feature and each list claims "this is not face" or "this could be face." It can be used to determine whether a portion of an image is a face. It provides a face detector when combined with the sliding window algorithm. This file is specifically for frontal faces only, and the "alt" indicates that it is not the standard OpenCV solution, but rather the result of a different—alternative—data set and training [7, 8].

haarcascade_lefteye_2splits

haarcascade_lefteye_2splits is an XML file containing a serialized Haar cascade detector for the left eye.

haarcascade_righteye_2splits

haarcascade_righteye_2splits is an XML file containing a serialized Haar cascade for the right eye [3].

DeepFace

DeepFace is Python's most lightweight face recognition and attribute analysis library. The open-source DeepFace library comprises all cutting-edge AI models for face recognition and executes all facial recognition tasks in the background automatically [9].

Setting up Deepface

Install the DeepFace package

pip install deepface

Import the Library in Python

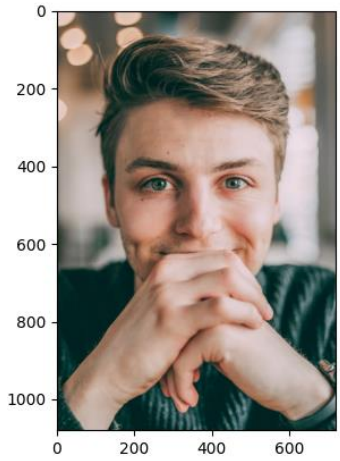

Figure 1. Neutral person.

```
Python 3.9.10 (tags/v3.9.10:f2f3f53, Jan 17 2022, 15:14:21) [MSC v.1929 64 bit (
AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
== RESTART: C:\Users\Zaheen\Desktop\Sem III Mini Project\Research\myEmotion.py =
1/1 [==============================] - ETA: 0s
1/1 [==============================] - 0s 251ms/step
Angry:  0.0013860965736967046
Disgust:  4.0210403540186235e-05
Fear:  0.015076552517712116
Happy:  34.781548380851746
Sad:  0.5437877029180527
Surprise:  0.0008905143658921588
Neutral:  64.65727090835571
>>>
```

Figure 2. Result of analysis.

Speech and Visual Speech Recognition

Figures 3 and 4 Active speaker detection (ASD) is the task of identifying the speaker in a visual scenario with one or more speakers. To achieve accurate ASD, it is crucial to interpret auditory and visual information both in the short and long term, and to account for audio–visual interactions. Previous approaches have relied solely on short-term data for making instant decisions, but TalkNet takes a unique approach by incorporating both short-term and long-term aspects (Figures 5 and 6). It utilizes self-attention mechanisms to capture long-term speaking patterns, audio and visual temporal encoders for feature representation, and audio–visual cross-attention mechanisms for inter-modal interaction. TalkNet has been tested on the AVA-ActiveSpeaker data set and the Columbia ASD data set, showing an improvement over state-of-the-art systems by 3.5% and 2.2%, respectively [2].

Figure 3. Happy person.

```
== RESTART: C:\Users\Zaheen\Desktop\Sem III Mini Project\Research\myEmotion.py =
1/1 [==============================] - ETA: 0s
1/1 [==============================] - 0s 94ms/step
Angry:  1.2641326456717722e-10
Disgust:  1.1822128219366278e-19
Fear:  4.933122719944914e-13
Happy:  99.99105334228595
Sad:  8.09126238213279e-12
Surprise:  1.476089116766367e-07
Neutral:  0.00894957537322568
>>>
```

Figure 4. Result analysis.

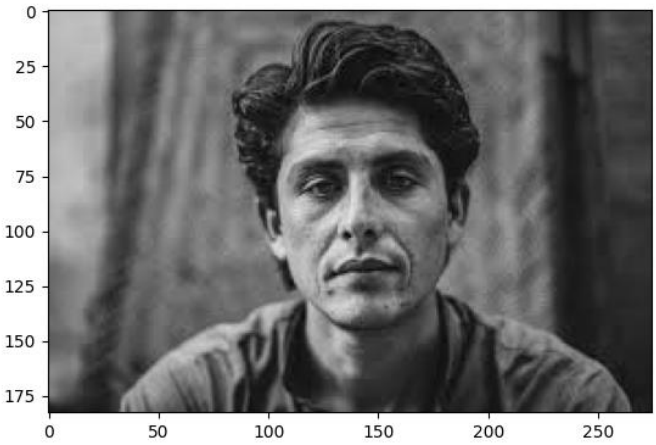

Figure 5. Sad person.


```
== RESTART: C:\Users\Zaheen\Desktop\Sem III Mini Project\Research\myEmotion.py =
1/1 [==============================] - ETA: 0s
1/1 [==============================] - 0s 85ms/step
Angry:  9.492392092943192
Disgust:  2.4703368239897827e-06
Fear:  0.4609024152159691
Happy:  0.04070608119945973
Sad:  54.91060018539429
Surprise:  0.0005346216767065926
Neutral:  35.09485423564911
>>>
```

Figure 6. Result analysis.

Yawning Module

Input models: Haarcascade models: Github

Figure 7. Sleeping module under test.

Detecting Facial Movement

The proposed system uses the following different components for attention level change detection:

- Facial landmark detection (LV)
- Emotional AI attribute model (ML)
- Artificial Emotional Intelligent Behavior model
- Monitor behavior continuously in video frame

One of the essential characteristics for detecting attention decreases is now head posture (Figure 7). Changes in attention level are directly correlated with head-turning posture. An individual learner is paying attention if their head is up or facing forward. The student is likely distracted if the head is spinning, translating, or tilting back [10].

ANALYSIS AND FINDINGS

A Google form was circulated among a crowd, and the response was taken anonymously, the identities of these students were not stored as if the details were recorded, the accurate result would have never been accurate (Figures 8–11).

LIMITATIONS AND FUTURE SCOPE

This article aims to capture emotions from students; however, the emotion spectrum of every student is not going to be the same, it will be different.

The system might recognize a student as *interested* but the student can mimic the idea of paying attention and vice versa. This system is prone to provide misleading results, however, if at certain a particular class is studied based on the emotions reflected by a particular student then the system will provide accurate results (Figures 12–15).

The system (*proposed*) might generate false results in cases where the emotion's completely fall out of the known emotion list.

The system (*proposed*) will be incompatible with the lighting of a particular class, to add on to the same point, a single central camera possibly will not capture the entire class, thus cameras should be set up from different angles.

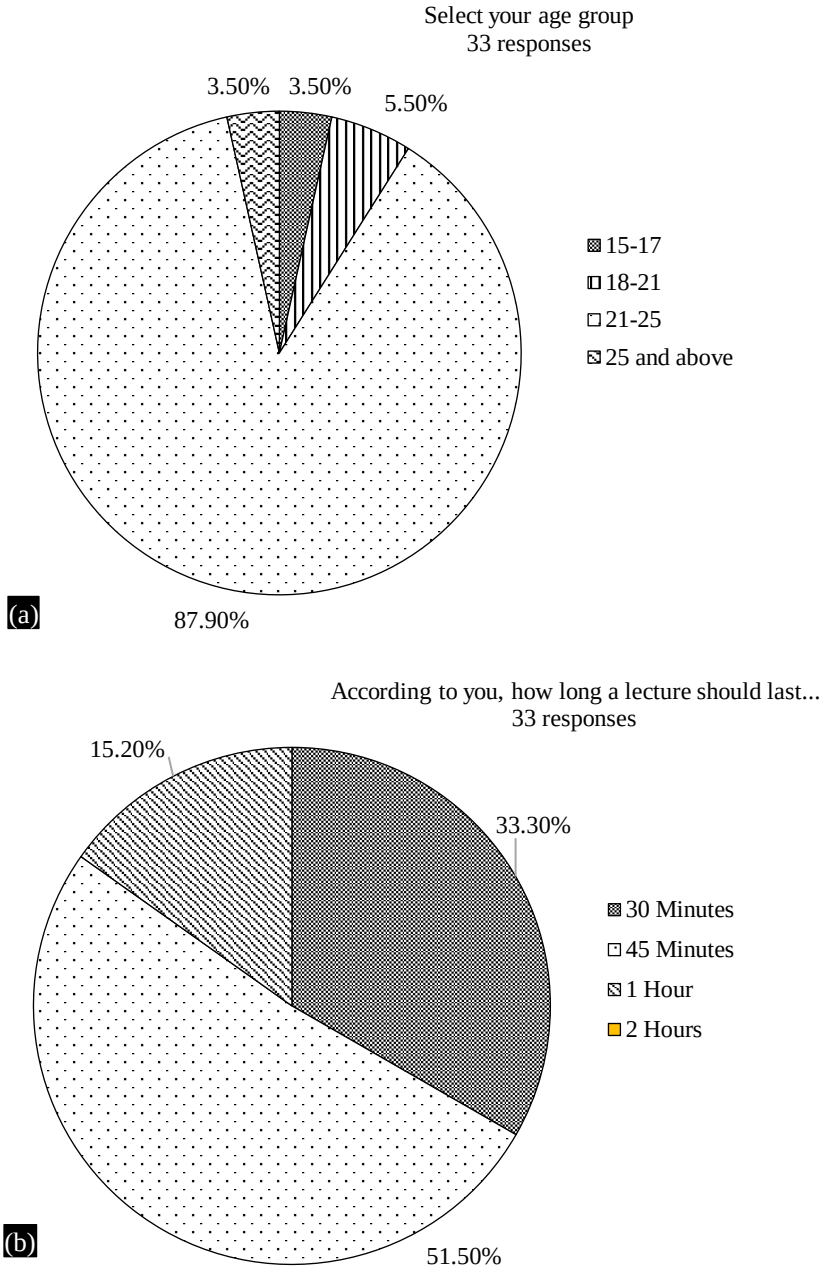

Figure 8. (a) The majority of the form responses were from the age group 21 to 25 years and (b). Views on lecture duration.

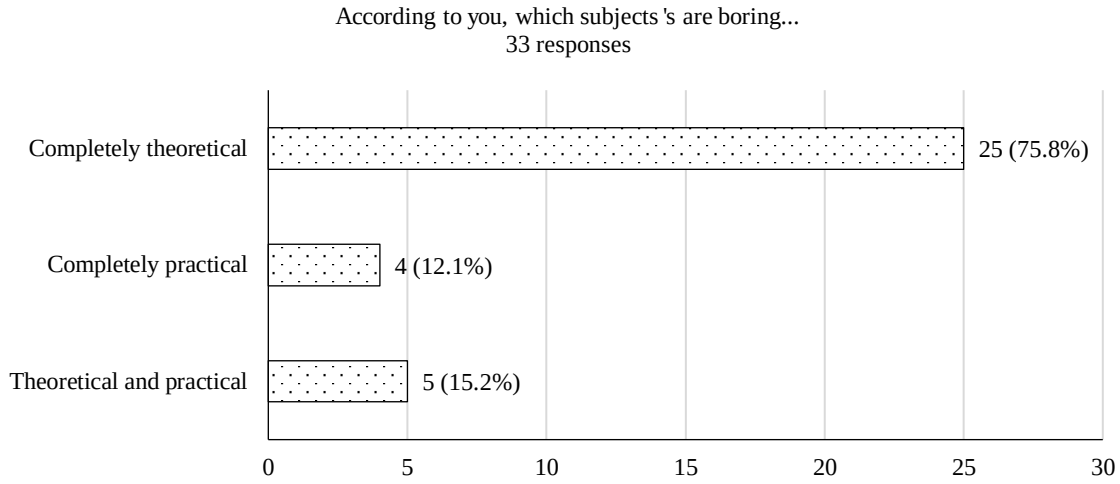

Figure 9. Views on subject choice.

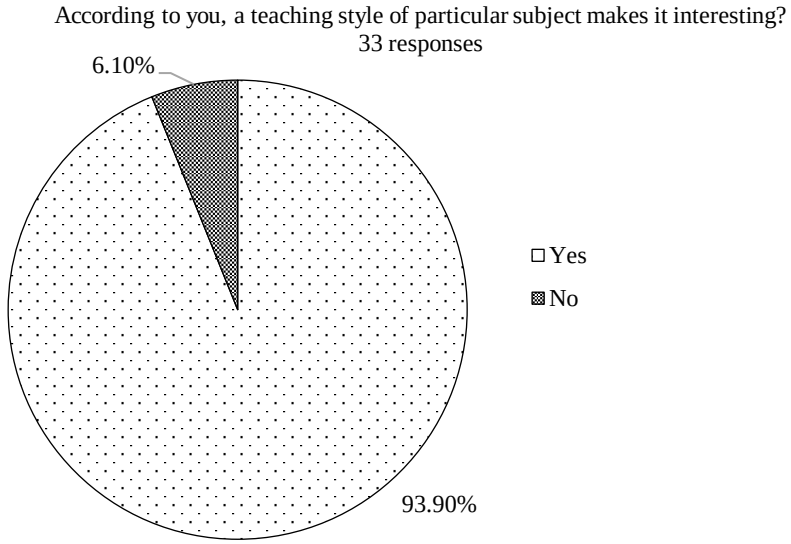

Figure 10. Views on boring subject.

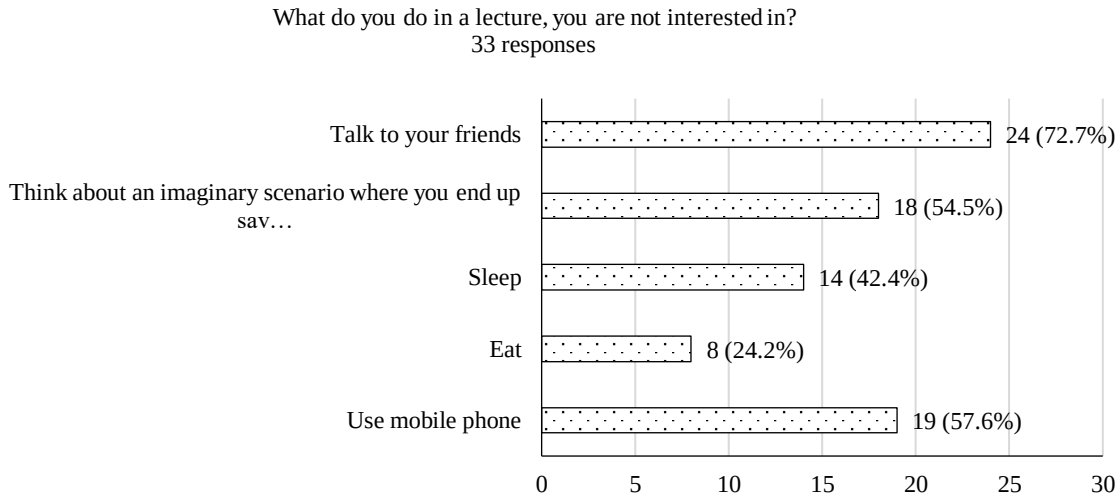

Figure 11. If a student is not interested, he/she/they might do these activities.

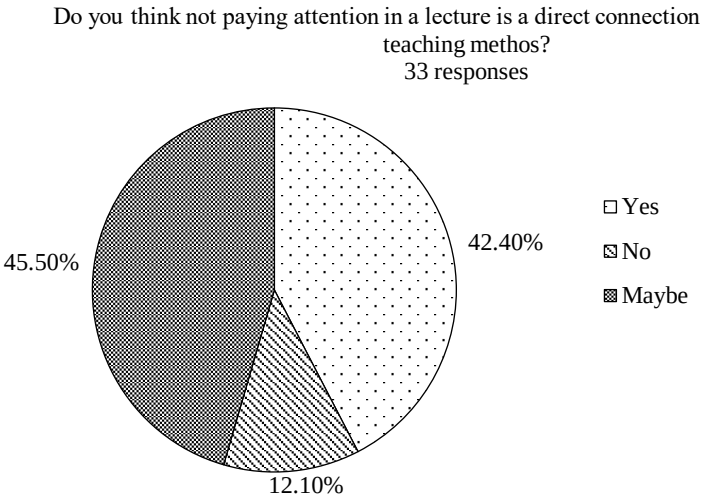

Figure 12. Views on teaching style.

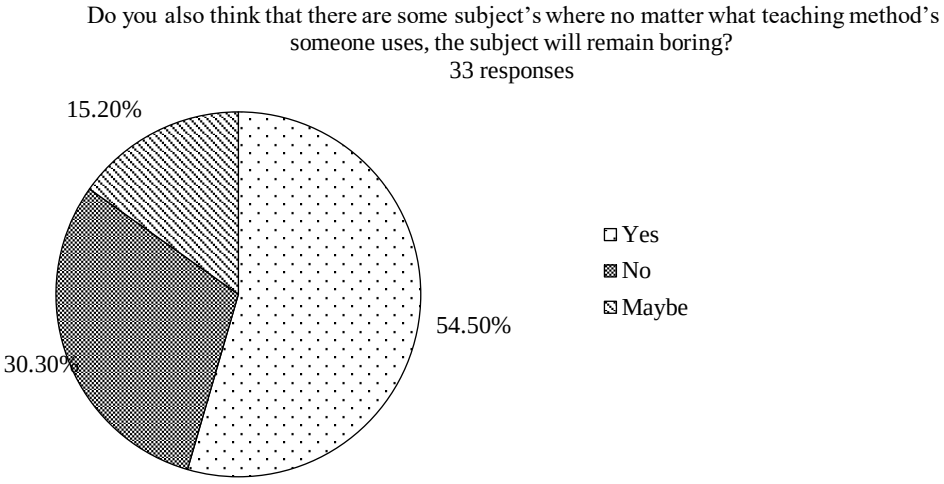

Figure 13. Views on Teaching style and a boring subject.

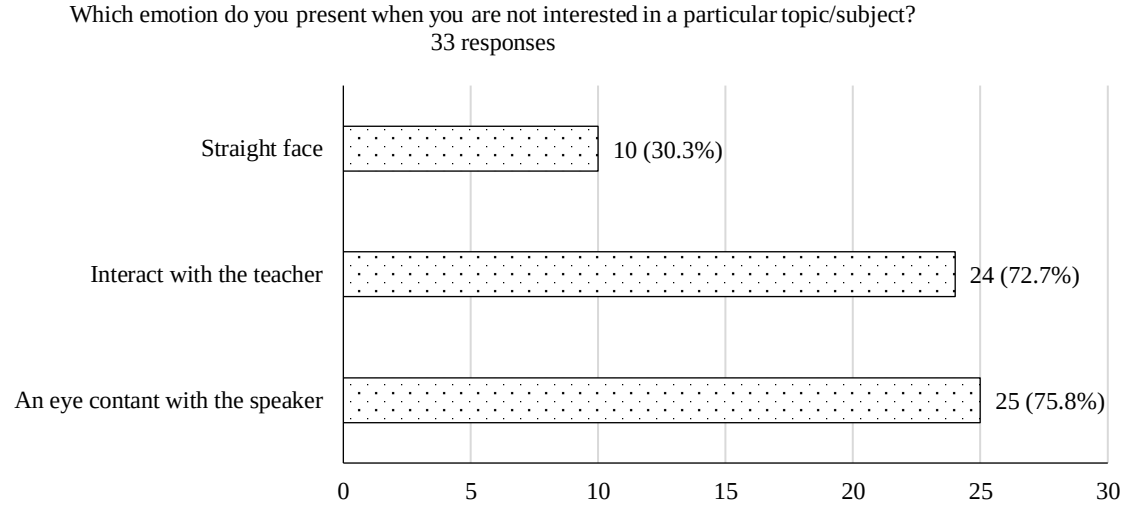

Figure 14. Views regarding subject interest.

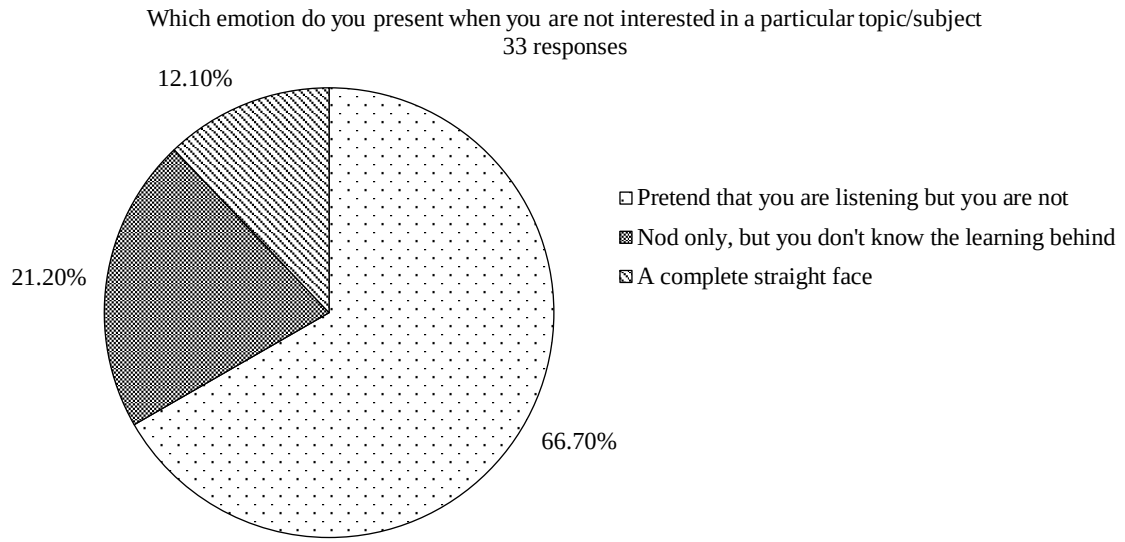

Figure 15. Views on the interest of the particular subject.

Through the above research, it is clear that not every student is going to be interested in a lecture. Most of the student's favor completely theoretical subjects, but in reality, practical subjects mark a lesson which stays with the student.

If a student is not interested in a subject, the student will not be interested in that subject during the entire course. Teaching styles do have an impact, but certainly a student changing their mind state is entirely off board.

While designing this system, the system should embed certain emotions clearly, as the human emotions are mostly unpredictable. Designing such system, is a breathtaking task as student's sometimes reflect a *straight face* emotion even if interested or not interested.

The system (*proposed*) can generate falsified results in some cases, where the same set of emotions are depicted by students, that is, A Straight Face.

CONCLUSION

All the modules like Face Recognition Module, Drowsiness Detection Module, Head Detection Module should be combined in one single entity, thus delivering the precise functionality.

The Student Emotion Recognition System (*proposed*) will help filter teachers based on teaching skills. Not every teacher is going to make a subject interesting, but this can be improved by analyzing the student's attention span in that lecture.

REFERENCES

1. Jaiswal A. (2022). Face recognition system using python. [Online] Analytics Vidhya. Available at https://www.analyticsvidhya.com/blog/2022/04/face-recognition-system-using-python/ [Accessed on April 2023]
2. Tao R, Pan Z, Das RK, Qian X, Shou MZ, Li H. Is someone speaking? Exploring long-term temporal features for audio-visual active speaker detection. In: Proceedings of the 29th ACM international conference on multimedia; 2021 Oct 17: 3927–3935 pp. doi: 10.1145/3474085.3475587.
3. 3Anaustinbeing. (2018). GitHub: anaustinbeing/haar-cascade-files. Every haar-cascades here! [Online] GIThub. Available at https://github.com/anaustinbeing/haar-cascade-files [Accessed on April 2023]

4. De Silva LC, Miyasato T, Nakatsu R. Facial emotion recognition using multi-modal information. In: Proceedings of the ICICS International Conference on Information, Communications and Signal Processing; Vol. 1. Singapore: IEEE Publications; 1997. doi: 10.1109/ICICS.1997.647126.
5. Harms MB, Martin A, Wallace GL. Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. Neuropsychol Rev. 2010; 20 (3): 290–322. doi: 10.1007/s11065-010-9138-6, PMID 20809200.
6. Jain DK, Shamsolmoali P, Sehdev P. Extended deep neural network for facial emotion recognition. Pattern Recognit Lett. 2019; 120: 69–74. doi: 10.1016/j.patrec.2019.01.008.
7. Turetsky BI, Kohler CG, Indersmitten T, Bhati MT, Charbonnier D, Gur RC. Facial emotion recognition in schizophrenia: when and why does it go awry? Schizophr Res. 2007; 94 (1–3): 253–263. doi: 10.1016/j.schres.2007.05.001, PMID 17583481.
8. Feinberg TE, Rifkin A, Schaffer C, Walker E. Facial discrimination and emotional recognition in schizophrenia and affective disorders. Arch Gen Psychiatry. 1986; 43 (3): 276–279. doi: 10.1001/archpsyc.1986.01800030094010, PMID 3954548.
9. Serengil S. (2023). DeepFace: the most popular open source facial recognition library. [Online] viso.ai. Available at https://viso.ai/computer-vision/deepface/ [Accessed on April 2023]
10. Sharma G. (2020). Facial emotion recognition using convolutional bidirectional LSTM medium. [Online] Analytics Vidhya. Available at https://medium.com/analytics-vidhya/facial-emotion-recognition-using-convolutional-bidirectional-lstm-49d4199024f [Accessed on April 2023]

