SHAASTRA2024

AT Makeathon

The Amazing Team

Cambridge Institute of Technology, North Campus
Department of Computer Science & Engineering

Sairaj Todkar
(Leader)
Stream: Computer Science
Year of graduation:2026

Srijani M
Stream: Cyber Security
Year of graduation:2026

Rahul Raj
Stream: Computer Science
Year of graduation:2026

Problem Statement #3

Assistive Learning Tools

Develop an educational aid that supports students with diverse learning needs (dyslexia, ADHD, etc.) to enhance etc.) to enhance their learning experiences and academic achievements in classrooms.

Deep into the problem statement

The loss of attention during long online meetings, affecting productivity for employees and learning quality for students, which can directly result in fall of trust in online education and job performance (remote software job).

The trust on offline tuitions are always high compare to online tuitions due no proper attraction & attention of student on online classes and offline tuition are always cost-high and have many issues like :

- *Financial strain and time wastage experienced by students in traveling and living cost .*
- *Offline Tuition centers, where a significant portion of funds is invested in infrastructure for accommodating a large number of students, detracting from the primary focus on quality teaching and educator development.*

Proposed solution / Idea

1

Diverse Learning Needs

Our Solution: Explicit support for diverse needs, including dyslexia and ADHD. Evaluation: Features like dyslexia-friendly learning and ADHD support align with diverse learning needs.

2

Enhancing Learning Learning Experiences Experiences

Our Solution: The iCB, iCB, personalized examples, and interactive elements improve the learning learning experience. Evaluation: Innovative Innovative features foster active engagement and enhance the overall learning experience.

3

Academic Achievements

Our Solution: Continuous assessment, assessment, progress tracking, and adaptable adaptable teaching strategies. Evaluation: Evaluation: Features promote ongoing assessment, aiding in academic achievements. achievements.

4

Inclusivity and Accessibility

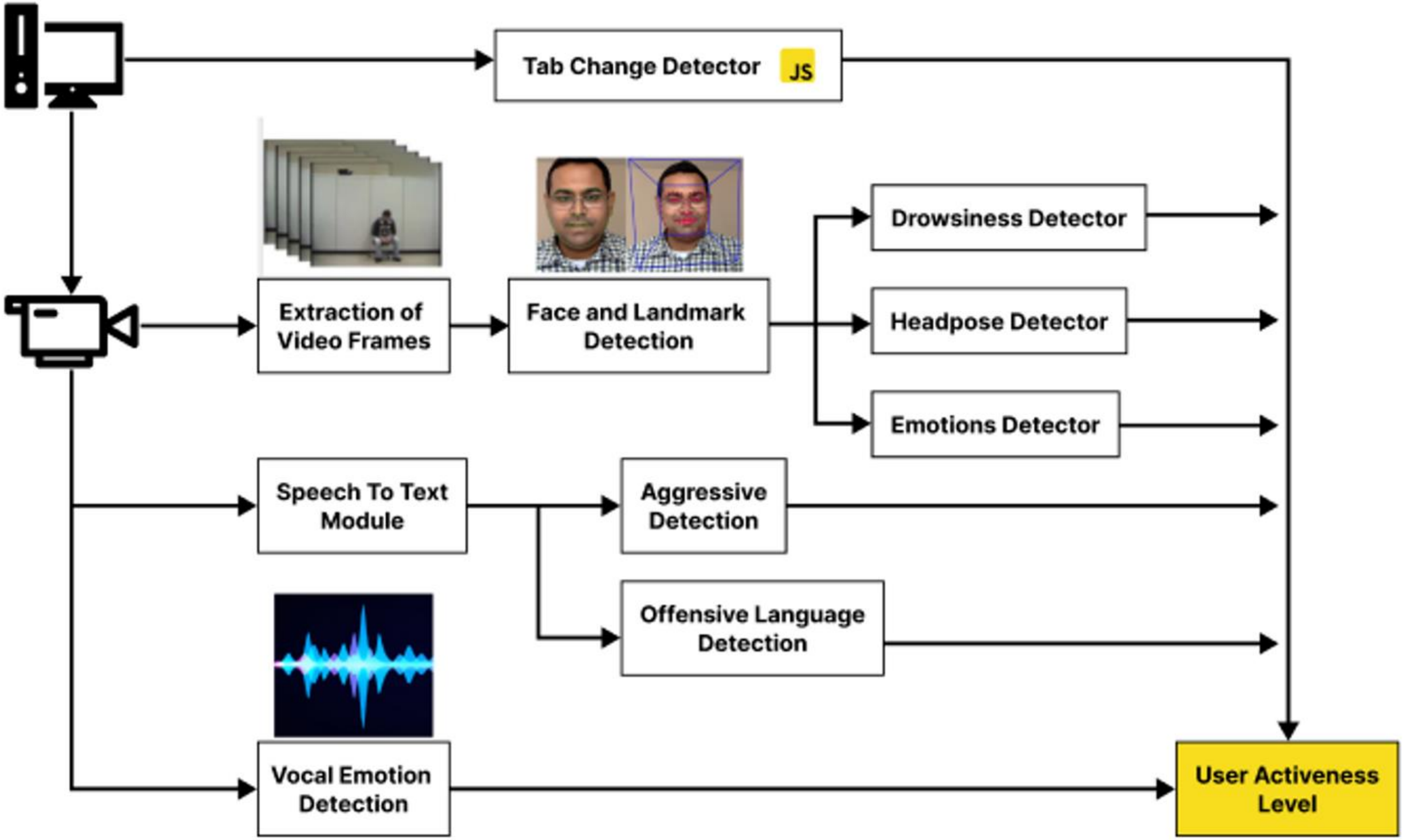
Our Solution: Accessibility Accessibility features for inclusivity. Evaluation: Designed with inclusivity in in mind, considering accessibility needs.

5

Cost-Effective Implementation and Offline Offline Environment

Our Solution: Emphasis on cost-effectiveness and creating an offline class class environment. Evaluation: Acknowledgment of financial constraints constraints and offline class feature contribute to practical and scalable implementation.

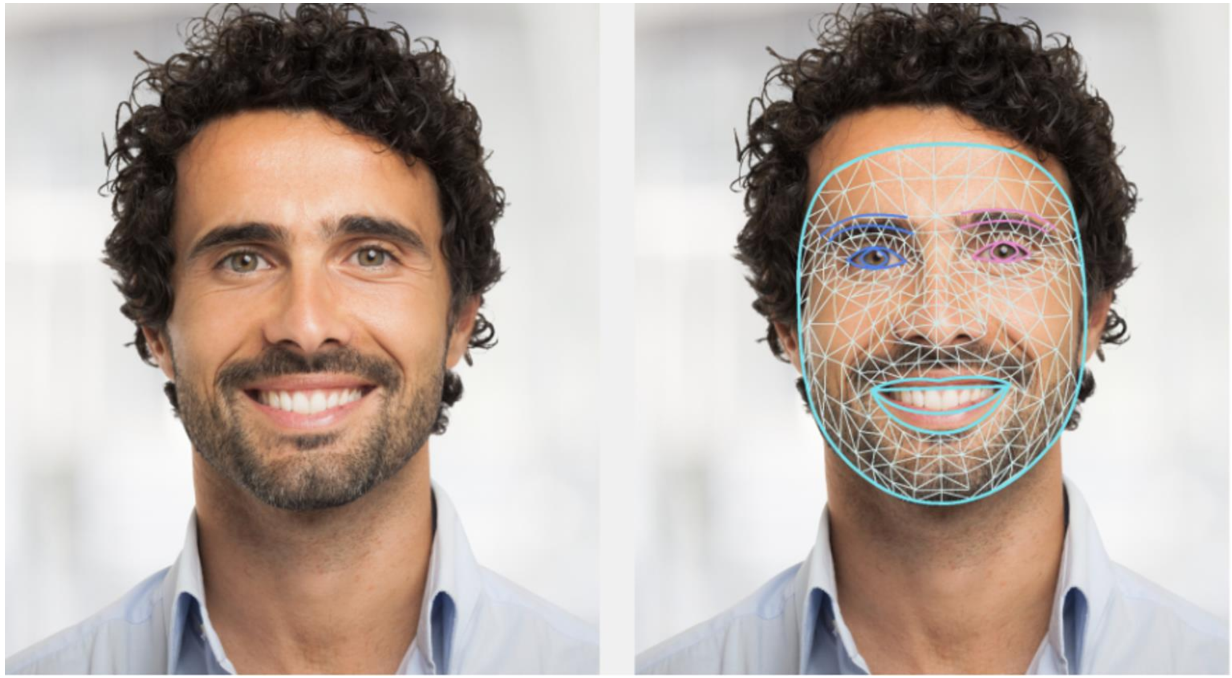
Flow Diagram



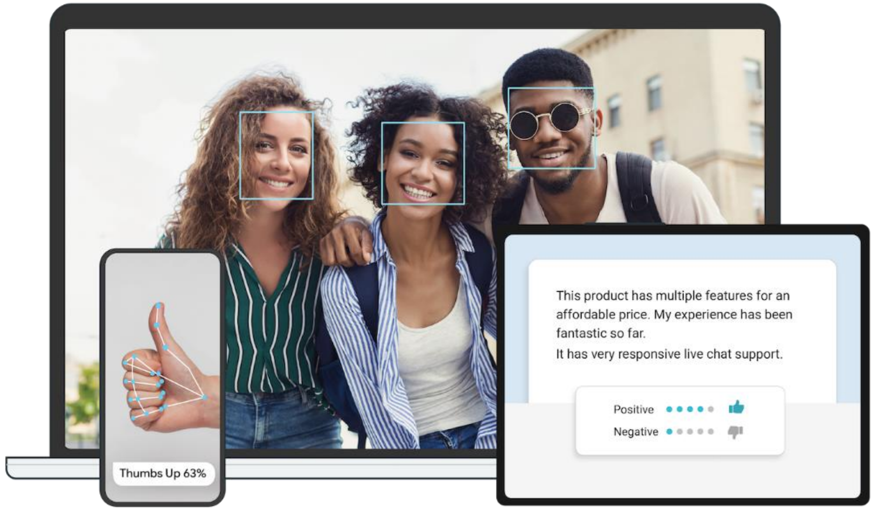
Attention Monitoring

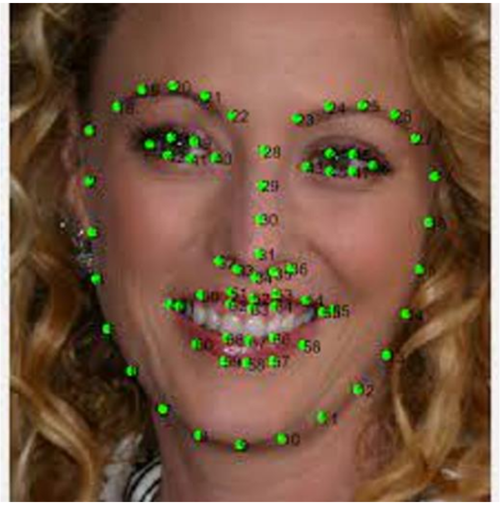
Face and Landmark Detection

MediaPipe Solutions provides a suite of libraries and tools for you to quickly apply artificial intelligence (AI) and machine learning (ML) techniques in your applications. You can plug these solutions into your applications immediately, customize them to your needs, and use them across multiple development platforms. MediaPipe Solutions is part of the MediaPipe open source project, so you can further customize the solutions code to meet your application needs.



Reference
https://github.com/google/mediapipe
https://developers.google.com/mediapipe





Attention Monitoring

Drowsiness Detection

The first step in the drowsiness detector is detecting the face, where a 68-point structure is distributed among the recognized key points in a human face. Then the drowsiness of the student will be calculated using the eye aspect ratio (EAR) and the yawn aspect ratio (YAR) (Shah et al., 2021). Figure 2 shows the presentation of the 68 facial landmarks (Pinzon-Gonzalez & Barba-Guaman, 2022). EAR and YAR are calculated by Eq. 1 and Eq. 4 for drowsiness detection. The videos will be processed using the OpenCV library (Howse, 2013). For each frame, the EAR and the YAR are extracted. Hence the EAR and YAR are calculated for each second for each student. A threshold for the EAR indicates when the eyes are closed, indicated as sleeping (Shah et al., 2021). A threshold for YAR indicates when the mouth is opened widely, indicated as yawning (Shah et al., 2021). The drowsiness detector generates the first two inputs that we have for our ML model.

The EAR is defined as (Shah et al., 2021)

$$EAR = \frac{EAR1 + EAR2}{2}$$

(1)

$$EAR1 = \frac{||37-41|| + ||38-40||}{2\,||36-39||}$$

(2)

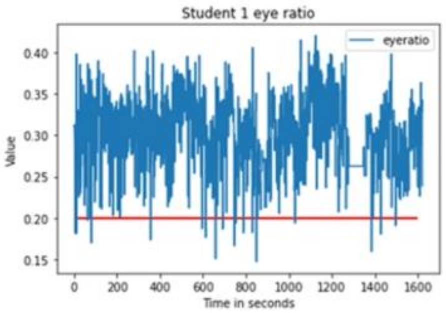
$$EAR2 = \frac{||43-47|| + ||44-46||}{2\,||42-45||}$$

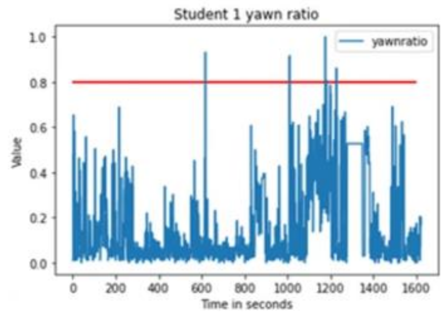
(3)

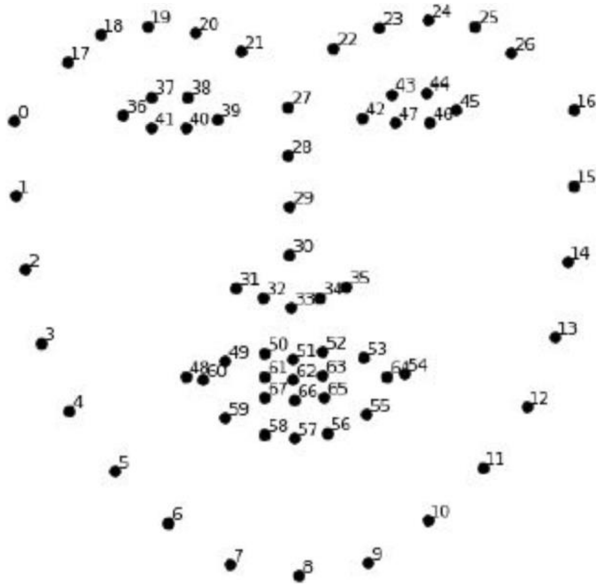
The YAR is defined as (Shah et al., 2021)

$$YAR = \frac{||61-67|| + ||62-66|| + ||63-65||}{||64-60||}$$

(4)







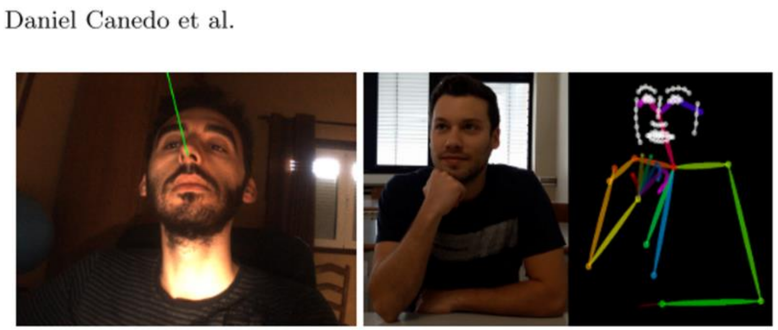
Reference
https://www.semanticscholar.org/paper/Assessment-of-Student-Attentiveness-to-E-Learning-Shah-Meenakshi/effba10a9dcec165e201fd55c77a6799e4f1daed

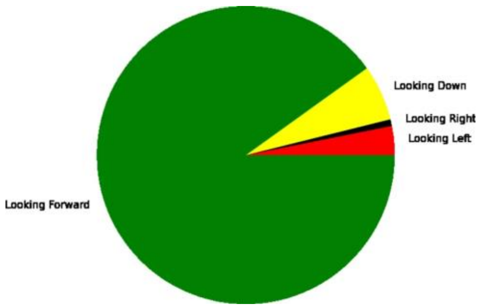
Attention Monitoring

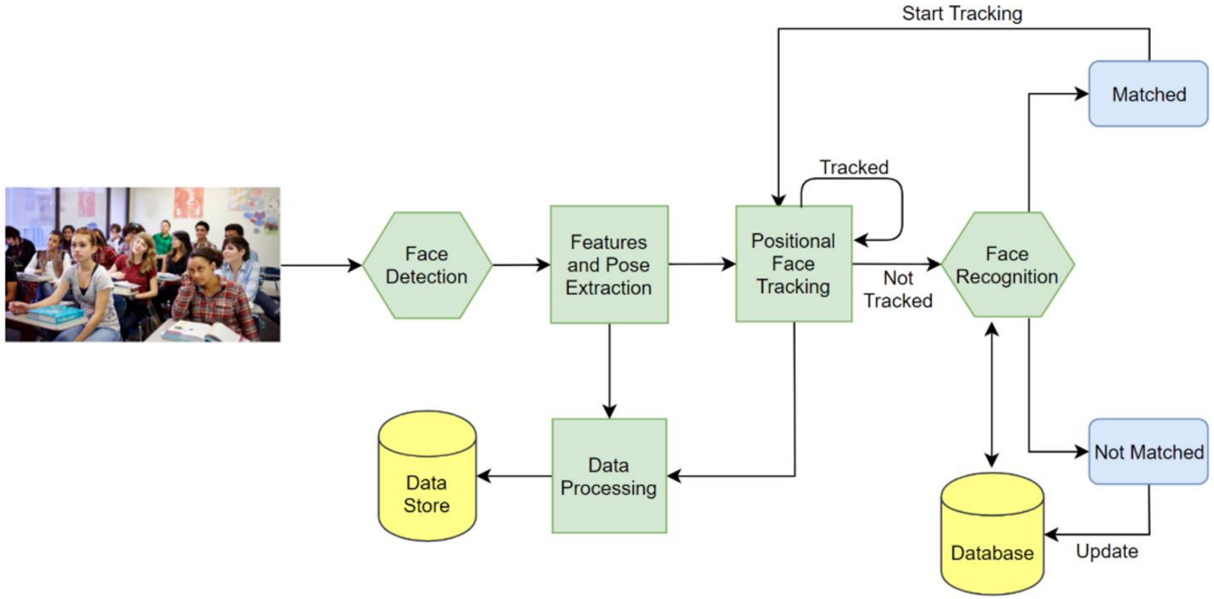
Headpose Detector

The head pose can help show the student's distraction or attentiveness (Pinzon-Gonzalez & Barba-Guaman, 2022). When a student is distracted, they may start looking here and there (Shah et al., 2021). Hence, the head's position may help recognize students' attentiveness and assist in training the ML model. First, a face mesh is built to identify the face and its six key points. Then the rotation angle is calculated. The X and Y components of the rotation angle are determined using the OpenCV library (Howse, 2013). For each frame, the two components of the rotation angle are extracted. Hence the two components of the rotation angle are calculated for each second for each student generating two more inputs for our ML model.

We propose the use of the MTCNN. This block is responsible to feed the Features and Pose Extraction block with images of the students' faces, which has the role of retrieving their facial features and body pose. We propose the adjustment of the head pose estimation, the head orientation and the use of OpenPose for the body pose







Reference

Pinzon-Gonzalez, J. G., & Barba-Guaman, L. (2022). Use of Head Position Estimation for Attention Level Detection in Remote Classrooms. *In Proceedings of the Future Technologies Conference (FTC) 2021 1*, 275–293. Springer International Publishing.

Attention Monitoring

Emotions Detector

Several authors have requested combining emotions with other measurement forms to assess student attentiveness (Revadekar et al., 2020; Shah et al., 2021), as facial expressions are one of the most potent signals for human beings to transfer their emotional states (Li & Deng, 2020). Facial expression recognition (FER) has been used in various study types, including driver fatigue surveillance, student attentiveness, and medical treatment (Khaireddin & Chen, 2021; Li & Deng, 2020). FER has been used to encode expression representation from facial representations. One of the famously used datasets for FER is FER 2013(Goodfellow et al., 2013; Khaireddin & Chen, 2021; Li & Deng, 2020). FER2013 is considered a benchmark in comparing performance for emotion recognition (Khaireddin & Chen, 2021). In (Khaireddin & Chen, 2021), the authors used convolution neural networks (CNN) where they adopted VGGNet architecture. Khaireddin & Chen fine-tuned the hyperparameters and experimented with various optimization methods for the VGGNet, where their model achieved an accuracy of 73.28% on FER2013 without extra training data. The VGGNet consists of four convolutional stages and three fully connected layers.

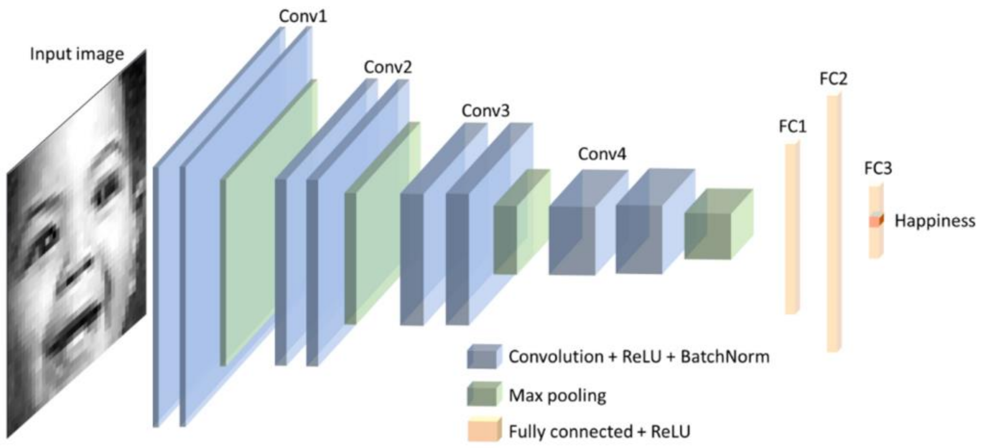
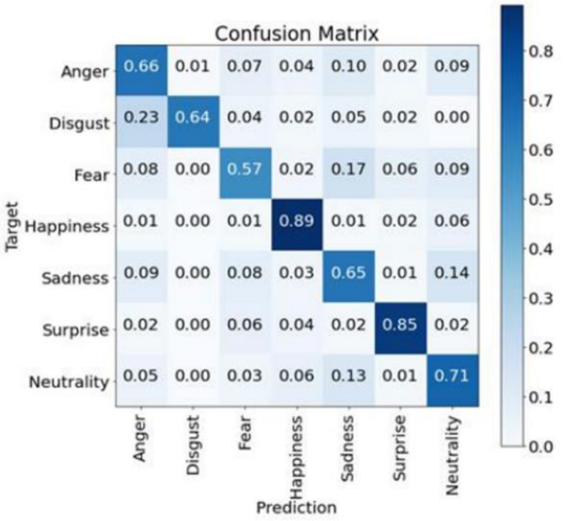


Figure 1 VGGNet architecture. A face expression image is fed into the model. The four convolutional blocks (Conv) extract high-level features of the image and the fully-connected (FC) layers classify the emotion of the image.



Method	Accuracy Rate
CNN [26]	62.44 %
GoogleNet [30]	65.20 %
VGG+SVM [29]	66.31 %
Conv + Inception layer [31]	66.40 %
Bag of Words [28]	67.40 %
Attentional ConvNet [27]	70.02 %
CNN + SVM [32]	71.20 %
ARM (ResNet-18) [33]	71.38 %
Inception [34]	71.60 %
ResNet [34]	72.40 %
VGG [34]	72.70 %
VGG (this work)	**73.28 %**

Reference

Khaireddin, Y., & Chen, Z. (2021). Facial emotion recognition: State of the art performance on FER2013. arXiv preprint arXiv:2105.03588

Attention Monitoring

	Aggression Detection			Offensive Language Detection		
PTLM	**Monolingual**	**Code mixed**	**Combined**	**Monolingual**	**Code mixed**	**Combined**
BERT$_{base}$	63.58±0.51	65.22±0.77	64.98±0.28	60.99±0.43	61.94±0.14	62.05±0.25
RoBERTa$_{base}$	**66.63±0.12**	65.42±0.61	62.13±0.89	**63.46±0.75**	62.06±0.48	60.21±0.30
XLM-R$_{base}$	65.49±0.73	66.85±0.22	**67.87±0.05**	61.24±0.31	64.42±0.02	65.41±0.73
HingRoBERTa	64.01±0.53	**66.94±0.53**	66.47±0.53	61.92±0.26	**64.97±0.13**	**65.45±0.21**
Bernice	63.49±0.15	61.13±0.43	62.75±0.82	60.88±0.57	59.01±0.38	60.58±0.16

Aggressive & Offensive Language Detection

We found that the results obtained via fine-tuning pretrained language models in this section. Table 4 reports the test set macro F1-scores from pre-trained language models for the two tasks of aggression detection and offensive language detection on our dataset. In addition to this, we also present the scores on English monolingual and Hindi-English code-mixed subsets of our dataset. For aggression, we observe that XLM-Rbase outperforms other pre-trained language models on our overall dataset, achieving the highest macro F1score of 67.87. On the English subset, we observe that RoBERTabase performs better than other models with a macro F1-score of 66.63, whereas for the Hindi-English code-mixed subset, Hing-RoBERTa gives the best macro F1-score of 66.94. For offensive language detection, we observe that Hing-RoBERTa outperforms other pre-trained language models on our overall dataset, achieving the highest macro F1-score of 65.45. On the English subset, we observe that RoBERTabase outperforms other models with a macro F1-score of 63.46. For the Hindi-English code-mixed subset, Hing-RoBERTa once again gives the best performance with a macro F1-score of 64.97.

Reference

Towards Safer Communities: Detecting Aggression and Offensive Language in Code-Mixed Tweets to Combat Cyberbullying

Nazia Nafis, Diptesh Kanojia, Naveen Saini, Rudra Murthy, Indian Institute of Information Technology Lucknow, India.

Attention Monitoring

Vocal Emotion Detection using Multilayer Perceptron

Emotions portrayed	Emotion judgements								Classification errors
	Anger	**Fear**	**Happiness**	**Neutral**	**Sadness**	**Disgust**	**Surprise**	**H_u**	**Listeners [$CI_{95\%}$]**
Anger	**96.07**	0.10	0.83	0.42	0.10	0.48	2.00	0.835	3.93% [3.25%; 4.70%]
Fear	0.52	**77.72**	0.97	2.93	13.93	0.41	3.52	0.731	22.28% [20.77%; 23.84%]
Happiness	0.56	0.03	**75.17**	12.59	0.10	0.17	11.38	0.397	24.83% [23.26%; 26.44%]
Neutral	1.03	0.73	1.38	**91.69**	3.14	0.62	1.41	0.684	8.31% [7.33%; 9.38%]
Sadness	0.45	3.24	0.83	3.97	**90.23**	0.62	0.66	0.752	9.76% [8.71%; 10.90%]
Disgust	8.52	0.69	5.11	10.28	0.72	**67.51**	7.17	0.652	32.49% [30.79%; 34.23%]
Surprise	3.41	0.17	58.14	0.97	0.07	0.10	**37.14**	0.218	62.86% [61.07%; 64.62%]
Total	-	-	-	-	-	-	-	-	23.50% [22.91%; 24.09%]

Vocal emotion detection using Multilayer Perceptron (MLP) neural networks has proven to be a powerful and effective approach to understanding and classifying emotions in spoken language. With the availability of large annotated datasets and advancements in deep learning techniques, MLP-based models continue to improve the accuracy and robustness of vocal emotion detection systems, contributing to enhanced human-computer interaction and emotional analysis applications.

Reference

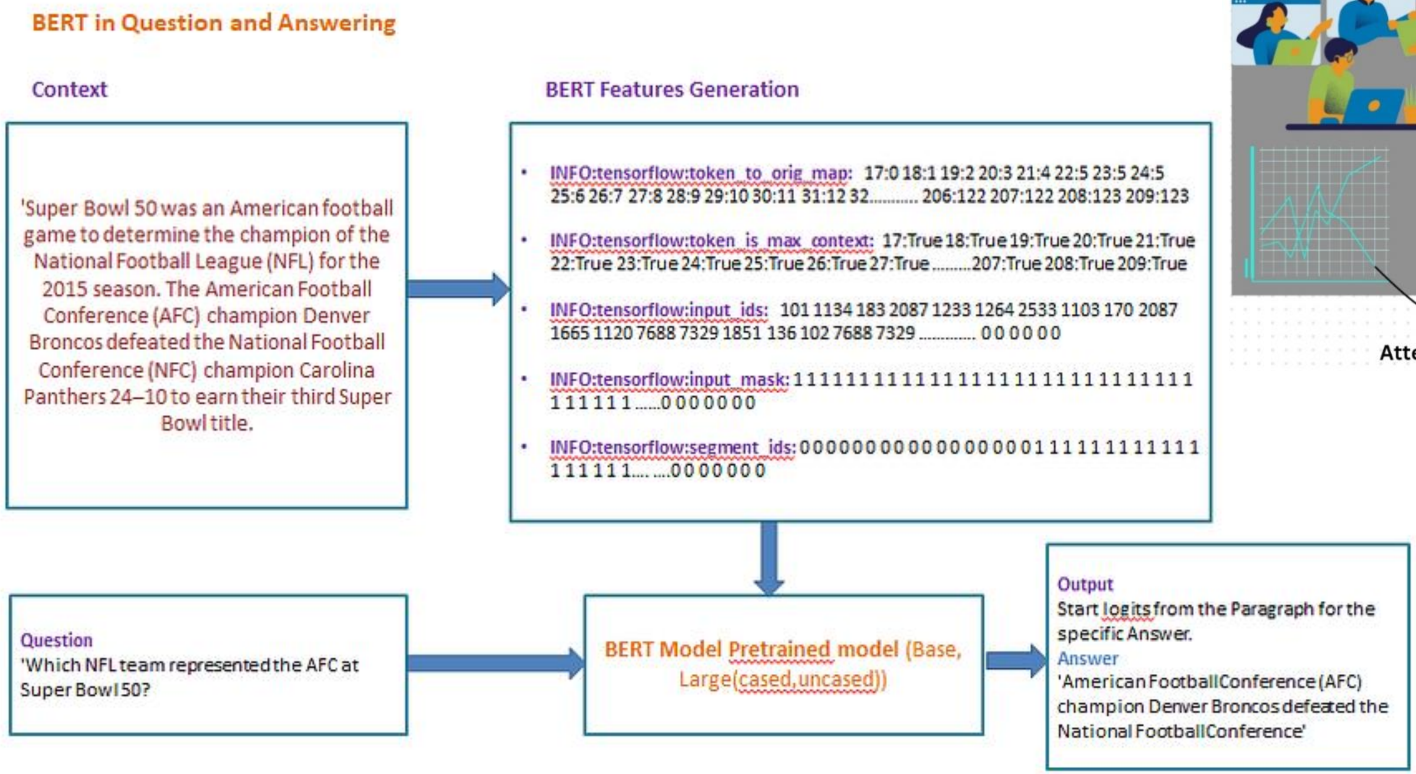
Lausen, A., Hammerschmidt, K. Emotion recognition and confidence ratings predicted by vocal stimulus type and prosodic parameters. *Humanit Soc Sci Commun* **7**, 2 (2020). https://doi.org/10.1057/s41599-020-0499-z

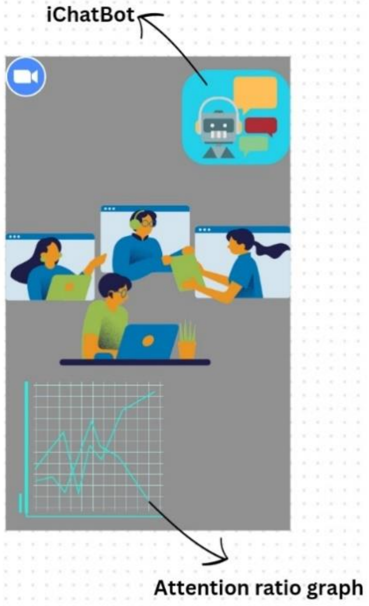
Continuous Assistant via Intelligent Chatbot

BERT-based Cross-Lingual Question Answering Chatbot (iChatBot)

iChatBot tries to ask questions during breaktime from the users to check their attentiveness during the class. The questions are self generated by the chatbot by listening to the question during class hours.

The Intelligent Chatbot (iChatbot) utilizes Natural Language Processing (NLP) for conversation, Information Retrieval to fetch internet data on the current topic, Machine Learning for cross-questioning to assess comprehension, and potentially Web Scraping and API Integration to access and provide up-to-date information from the web page link given by the host that allow the student/ one in the meeting to use it as browser without exiting the tab and grab more information on ongoing topic. And also create polls and crossquestion in online meet to make it more effective as offline meets .





Reference

https://medium.datadriveninvestor.com/extending-google-bert-as-question-and-answering-model-and-chatbot-e3e7b47b721a

How does our innovation accelerate change change with the power of Technology?

Streamlining Processes

Our innovation leverages the power of technology to technology to accelerate change by streamlining streamlining processes, enhancing efficiency, and and facilitating seamless communication.

Enhancing Efficiency

Through the integration of cutting-edge technological solutions, we empower businesses and businesses and individuals to adapt to the rapidly rapidly evolving landscape of the digital age.

How is our solution different/unique from other solutions in market.

1

Intelligent Chatbot (iCB)

Utilizes AI, NLP, and information retrieval for personalized, real-time assistance.

2

Multisensory Learning

Integrates visual, auditory, and kinesthetic elements for a holistic learning experience.
experience.

3

Real-time Well-being Monitoring

Landmark and drowsiness detection ensure proactive student intervention.

4

Global Connectivity

Enables real-time collaboration, breaking geographical boundaries in education.

5

Adaptability to Emerging Tech

Incorporates VR, AR, and cloud-based platforms for staying at the forefront of advancements.
advancements

Business Model & Future Plans

1. COVID-19 Acceleration:

1. The COVID-19 pandemic has accelerated the adoption of online learning.
2. It resulted in a significant increase in the demand for online learning platforms.

2. Personalized and Adaptive Learning:

1. A growing trend towards personalized and adaptive learning is observed.
2. Online learning platforms are incorporating AI and machine learning to offer customized learning experiences tailored to individual needs.

3. Increasing Competition:

1. There is a rising competition among online learning platforms.
2. This competition has led to an increase in the number of players in the market.

4. Market Growth:

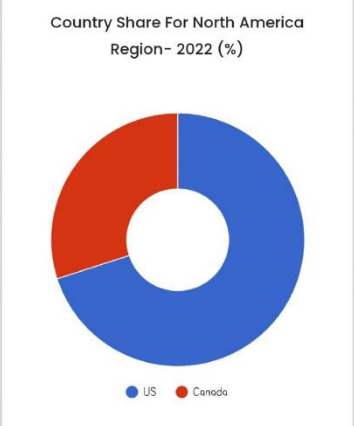
1. The global online learning platforms market is expected to reach $325 billion by 2025.
2. It is projected to grow at a compound annual growth rate of approximately 7%.

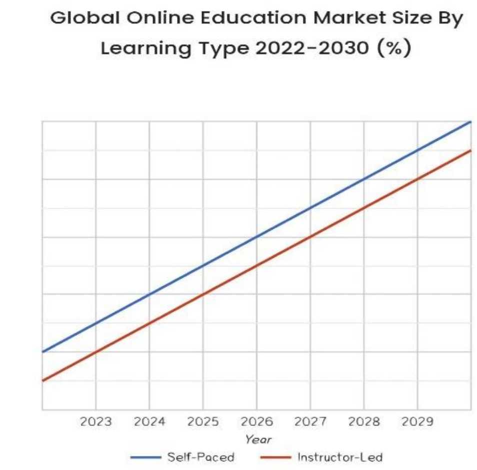
5. User Increase:

1. The number of users on online learning platforms has significantly increased in the past year.
2. This surge is attributed to individuals seeking flexible and accessible education options.

6. Future Growth Drivers:

1. The online learning platforms segment is expected to continue its growth trajectory.
2. It will be driven by increasing demand for remote and flexible learning options.
3. Advancements in technology will enable more personalized and engaging learning experiences





Thank you!