E-Learning App With Enhanced Functionalities To Assist Teachers And Students

Soujanyaah Karunanithy, Kishome Sivarajah, Viknatharshan Nadarajah, Sivashankar Sivaganeshamoorthy

Sri Lanka Institute of Information Technology, Faculty of Computing,
New Kandy Road, Malabe, Sri Lanka, PH- +94762613111
soujiakaru25@email.com

Sri Lanka Institute of Information Technology, Faculty of Computing,
New Kandy Road, Malabe, Sri Lanka, PH- +94773165543
kishosiva13@email.com

Sri Lanka Institute of Information Technology, Faculty of Computing,
New Kandy Road, Malabe, Sri Lanka, PH- +94703472492
docvik30@email.com

Sri Lanka Institute of Information Technology, Faculty of Computing,
New Kandy Road, Malabe, Sri Lanka, PH- +94776887978
hexshankar1997@email.com

Abstract: Since most institutions and schools are switching to online or hybrid techniques due to the COVID 19 pandemic, the use of e-learning has been expanding rapidly. The primary goal of this research study is to improve the effectiveness of the online educational system, thus we would be focusing on the technologies such as Video processing, Natural Language Processing and Artificial Neural Networks, to develop the web application and reduce the workload for instructors and students. Video Processing is one of the most accurate ways to monitor human behaviour. Eye-tracking is a technique for capturing real-time, objective user behaviour. Eye motions are quick, subconscious movements that communicate information that even the responder is unaware of. Converting audio of video lectures into text with noise removal using Natural Language Processing so students can view the lecture videos with subtitle and read the transcript after the class finished and also transcript will be translated into preferable language which was selected by the student for getting more clarification on lessons. From this, students can easily understand the lectures and save their time. A system also has the ability to work well with paragraphs of text as input. So that customized question papers can be prepared based on student feedback and evaluated for self-assessment automatically using Natural Language Processing. Chatbots have significantly reduced human effort by providing automated human-like solutions to a wide range of business and societal issues. In that list, this chat-bot which is developed using the Artificial Neural Network technology will answer the students' questions related to the subject and help them utilize their time in an effective manner with less effort while referring to a subject or preparing for the examination.

Keywords: Natural Language Processing, Video Processing, Artificial Neural Network, Web Application

1. Introduction

'Online learning' has become a much familiar term nowadays with almost every education provider switching to online platforms from physical classrooms due to the Covid19 pandemic. It is a method to deliver learning remotely to those who choose not to attend from a physical location. All around the world, educational institutions are looking online learning platforms to continue with the process of students. Online learning app is a technology-based study tool that allows information exchange that can be accessed from a smartphone. Students are getting more benefits through online learning. Teachers can deliver the lessons to students in efficient way through online learning platform and stud6ents can attend the classes from any location of their choice and they access the learning materials at a time of their comfort. Although the students get benefits through online learning, they face some problems with it. The main problem of the online classes has been the students getting distracted quickly and focus on other unwanted activities while the teacher is taking the class. It's not easy for the teacher to monitor each and every student during a class thus, some students take it as an advantage to do other fun-filled activities while pretending to be focusing on the classes. At the same time Most of the students are facing some problems like pace, intelligibility and network issues or even data shortage in online Learning platforms. The major problem is students are distracted by background noise while they are watching online lecture videos. Background noise is reduced the student's memory, motivation and learning ability. That's why I proposed "Automatic video transcription according to user-selected language with noise removal" which aims to reduce the spending time of students, increase the learning ability by noise removal and helps to get more clarification in the lessons by translate the transcript according to user selected language. And also, if the students can't attend the lectures on time, they can view recorded lectures with captions at any time. At the same time, they can download the transcript and read it any time to clear their doubts. On the other hand, Traditional teaching methods have been used since the early twentieth century and are still in use today. Most industries, particularly E-learning and education, have become digitalized as a result of technological advancements. Teachers can lessen the time-consuming problems by automating the question production and answer evaluation. Furthermore, according to the report, the majority of pupils had problems in various academic areas. As a result, we intend to create a system as a solution to this issue. Getting feedbacks from students following lectures, posing individualized questions, and grading the responses.

It will be quite beneficial for pupils to gain clarity on their problematic areas. Moreover, while referring to a subject, especially when the exams are near, student would have a limited amount of time, thus they would like to clear their doubts as soon as possible, that too without depending on another person. Even if they Google the questions, there is no guarantee that they would get the exact answers or definition that is required during the examination, since Google shows a vast variety of results. Also, it would be a time-consuming task if they opt to search through their notes or lecture slides. Therefore, the chat-bot that is going to be deployed on this application could help the students to use their time effectively. That's why we decided to implement this app for solving above mentioned problems. This paper is divided into five sections, which are as follows. Section 2 explains related work, section 3 offers a detailed explanation of our system, section 4 reviews the performance of our system and, section 5 summaries our work and concludes by discussing the scope for future work.

2. Literature Survey

Many methods were explored to improve the accuracy of the drowsiness detecting system. A model to identify tiredness based on electroencephalography (EEG) data has been put out by Mardi et al., [1]. For the purpose of discriminating between drowsiness and alertness, retrieved chaotic characteristics and the signal's logarithm of energy are extracted. With an accuracy rate of 83.3%, categorization was performed using an artificial neural network. A model to detect drowsiness based on Quality Signals, EEG, and Electrooculography was proposed by Noori et al. in [2]. A class separability feature selection approach was utilized to choose the optimal subset of characteristics. Using a self-organized map network, classification accuracy was 76.51 3.43%. Krajewski et al. [3] created a model based on steering patterns to detect tiredness. By utilizing cutting-edge signal processing techniques, three feature sets were created in this model to capture the steering patterns. Performance is assessed using five machine learning techniques, including SVM and K-Nearest Neighbor, and drowsiness detection accuracy was 86%. Yawning may be used to identify sleepiness, according to a method provided by Danisman, Abtahi, et al. This approach initially identified facial features before identifying mouth and eye locations. They computed a hole in the mouth due to the broad mouth opening. The face with the biggest gape is yawning. It is widely acknowledged that automatic captioning has the potential to improve teaching and learning for children with special educational needs, including those who speak a second language (e.g. Collins, 2013). Even if the difficulty of the key words in the captions is beyond the student's reading ability, the understanding of the video's content by the students can be improved (Ruan, 2015). According to Shiver and Wolfe's (2015) research of deaf students, many of them preferred watching videos with automatic captioning over those without it. Parton (2016) investigated how captions are used in connection to hearing pupils. She emphasizes the legal requirements that (higher) education institutions have to make materials accessible by pointing out the role subtitles play in enhancing accessibility of video resources [4]. In 2018, Kaustubh Bhattacharyya proposed employing several linear filtering algorithms to remove speech background noise [5]. Using an MFCC as a Feature Extraction with Frequency Sub-band Decomposition Technique, I. Patel and Dr. Y.S. Rao created an effective Speech Recognition system in 2010 [6]. When compared to MFCC without Sub-band Decomposition, these updated MFCC performs more accurately. By placing greater emphasis on the step of signal pre-processing, this system can operate more accurately [7]. Geeta Nijhawan and colleagues developed a real-time Hindi voice recognition system in 2014. Utilizing the Quantization Linde, Buzo, and Gray (VQLBG) approach, features were extracted using MFCC. To break the stillness, the Voice Activity Detector (VAC) was suggested [8]. An essential part of educational evaluation is automatic question generation (AQG). It takes a lot of time, money, and effort to create questions by hand and to manually evaluate answers. Researchers have been considering developing an autonomous system for generating questions and analyzing student responses for the past 20 years (Divate and Salgaonkar 2017) [9]. The procedure of cleaning and preparing the text for categorization is known as pre-processing the data. Online writings typically have a lot of noise and useless components like HTML tags, scripts, and ads. In addition, many of the words in the text have little bearing on the text's overall direction. Keeping those words increases the problem's high dimensionality and makes categorization more challenging because each word in the text is regarded as one dimension. The following is the premise of having the data appropriately pre-processed: lowering the text's noise level should help the classifier perform better and classify data more quickly, enabling real-time sentiment analysis. Online text cleaning, removing white space, extending abbreviations, stemming, removing stop words, managing negations, and feature selection are all phases in the process. The remaining phases are referred to as transformations, while the last step—which applies various functions to choose the necessary patterns—is referred to as filtering [10]. The words, concepts, or phrases that powerfully indicate an opinion as good or negative are known as features in the context of opinion mining. This indicates that they have a greater influence on the text's orientation than other words do. Using optical character recognition, text from PDF files is extracted in [11] using Apache PDFBox (OCR). Using the Over generating Transformations and Ranking method, question-answer pairs are formed from the resultant text document. In order to retrieve the response, the user input is compared to the question-answer pairings using pattern matching. [12] created a system that uses the provided text input to produce a variety of logical questions. The system employs a three-step strategy, in which it chooses the best possible set of sentences from the input text from which it could generate questions, looks for the sentence's subject and context to determine its main idea (Gap Selection), and then examines the best type of question that could be generated from that sentence (Question Formation).

In existing systems Manually Uploading Text Documents, Text Summarization, Answer Evaluation and Simplified User Interface are the common features in all applications. In this system additional and novel feature is generating personalized questions based on students' difficult areas. Several research have been conducted on chat-bot technology in which, significant amount of research have been on using the chat-bot system for educational

industry including chat-bots used, to provide administrative services (Hien et al., 2018, pp. 69–76; Rohrig & Heβ, 2019), to help assess students' performances (Benotti et al., 2017; Durall & Kapros, 2020, pp. 13–24), to learn and understand Computer Programming concepts (Okonkwo & Ade-Ibijola, 2020; Pham et al., 2018; Zhao et al., 2020), to answer students' queries (Clarizia et al., 2018, pp. 291–302; Ranoliya et al., 2017; Sinha et al., 2020, pp. 55–60). According to 2014 study, the Freudbot talkbot was created using ELIZA-style control capabilities and open-source software known as AIML (Artificial Intelligence Markup Language) [13]. The distinguishing quality of Freudbot was that it interacted with the students like a well-known historical figure. Even though the results were inconclusive, it was seen to have more promise for online education in the future [14]. Moreover, in their evaluation of the relevant literature on chatbots, Cunningham-Nelson et al. (2019, pp. 299–306) presented two situations in which chatbots might be utilized in a classroom and a prototype application for each. Smutny and Schreiberova (2020) employed a screening method in an impartial site directory to evaluate chatbots used for Facebook Messenger to support learning. Thomas (2020) looked at earlier research where chatbots helped both students and teachers, proving that the advantages outweigh the disadvantages and provide a more effective education [15]. Pérez et al. (2020) used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology to present an evaluation of various chatbots used for educational purposes [16].

3. Problem Definition

Students getting distracted quickly and focusing on other unwanted activities while the teacher is taking the class. It's too difficult for the teacher to monitor each and every student during a class. Most of the students are facing some problems like pace, intelligibility and network issues or even data shortage during online classes. The major problem is students are getting distracted by background noise while they are watching online lecture videos and the noise reduces the student's memory, motivation and learning ability. Every student is facing difficulties in different sections while attending the lectures or while watching the video lectures. Also, it's too difficult for the teacher to set questions manually and it may take a long time to proceed with it. While preparing for the exams or doing practice question, students might need to watch the lecture video or go through the lecture notes to get answers to the questions they are struggling. Incase if the student does not recognize under which section the question is falling, they might have a hard time finding appropriate answers.

4. Methodology

The system consists of four main components including the video processing unit that detects students who are getting distracted or falling asleep, automatic video transcription module with noise removal and translation facilities, automated personalized question generation and answer evaluation module and a chat-bot that uses neural network technology to answer students' questions. The initial step in developing the video processing component was to When faces are detected in photos, the Viola-Jones face detection technique is utilized as input to the Viola-Jones eye detection algorithm. The Violajones eye detection technique is used to extract the eye area from the facial pictures once the face has been identified, and CNN is then provided the input. The deep features are extracted using CNN with four convolutional layers, which are then fed to the fully connected layer. CNN then divides the photographs into sleepy and non-sleepy images using its Soft Max layer.

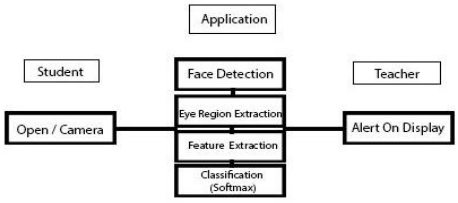

Figure 1: - High Level Architectural Diagram of student drowsiness detection System Diagram

Create automatic video transcription using NLP technology called "Speech Recognition". Speech recognition is a technology that enables machine to identity and understand words/phrases from spoken English and convert to machine readable format. NLP used to simplify speech recognition processes to make them less time consuming. Speech recognition and translation process steps:

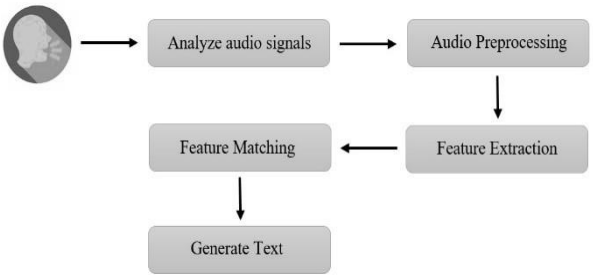

Figure 2: - Speech Recognition

Meanwhile, automatic video transcription system will extract the audio from the video recording file. Moviepy is a python library which is used to extract the audio. Then it will analyze the audio signals. Inside the audio signal analyze process, the system read the audio file, retrieving beats and visualize the audio signals. The first representation of audio signal is time-domain audio signal. It shows the amplitude of the sound wave with respect to time. Another representation of audio signal is frequency-Domain representation. The conversion of audio signal from its original time-domain state to frequency-domain state using Fourier Transform algorithm. By segmenting the sound wave into a spectrogram, we can apply the Fast Fourier Transform (FFT) technique to get a frequency representation of the sound wave. A spectrogram is a graphic depiction of a signal's frequency spectrum as it evolves over time. Then the system will detect the voice activity and filter the noise. It called audio preprocessing. It consists of cleaning the speech signal from ambient and undesirable noises, detecting speech activity, and normalizing the length of the vocal tract. In this study, noise is removed from noisy voice data in the frequency domain using spectral subtraction. With this technique, the noisy speech spectrum is calculated using the Fast Fourier Transform (FFT), and the average magnitude of the noise spectrum is then subtracted. [17] After the noise filtering system will extract the different features such as power, pitch and vocal tract configuration from audio signal using Mel Frequency Capstral Coefficients

(MFCCs). It is good in error reduction and able to produce a robust feature when the signal is affected by noise. The windowing of the signal, application of the DFT, calculation of the magnitude's log, warping of the frequencies on a Mel scale, and application of the inverse DCT are the main steps in the MFCC feature extraction approach. Then the system will automatically recognize the spoken words of a person based on given voice signals after comparing the extract features. Finally, it will generate the content as a text. User can read the transcript in paragraph format and download it. Then the user can translate the transcript. NLP technology is used to translate the text. The translation process contains two steps: language detection and language translation. In here system will detect the language of current data and translate the data to language of user choice.

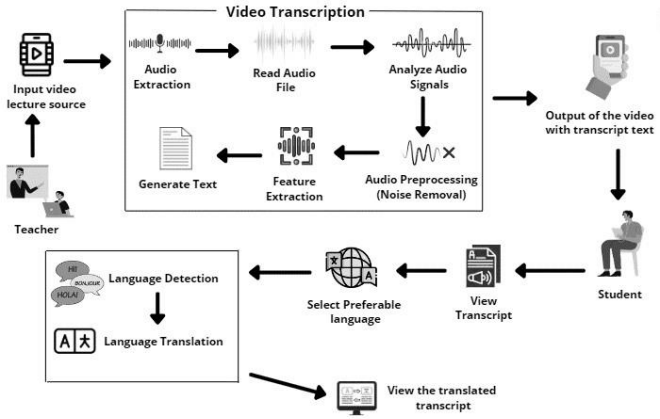

Figure 3: - System Overview Diagram Automatic Video Transcription According to User-selected Language with Noise Removal

The personalized question generation which is introduced in this system satisfy the student's requirement. Generating personalized questions based on students' difficult areas – After watching the lecture videos or attending online lectures students can give their feedbacks about the lecture. From the feedback using data pre-processing and text classification system will predict the students' difficult areas in a particular subject or module. By using T5 Transformer Model to generate questions with answers based the text that take as input based on the collected feedbacks. Also, by using some other models to generate different types of questions. Below list show the models and types of questions that generate in the system.

- Using T5 and word-net for generate MCQ Question Answering
- Using GPT-2 to generate True/False Question Answering

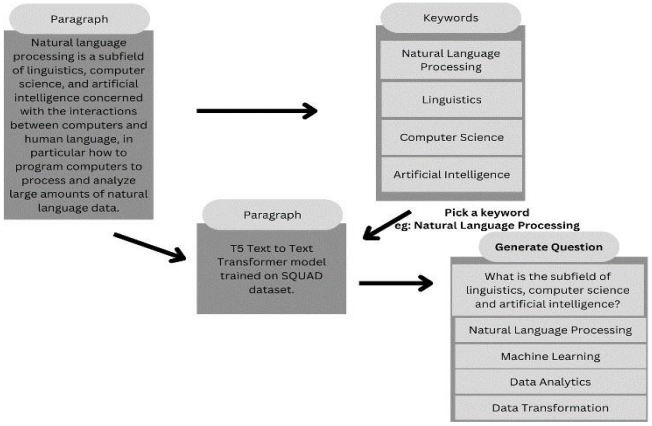

Figure 4: - Generating Question Based on the Reference Paragraph Using T5 Transfer

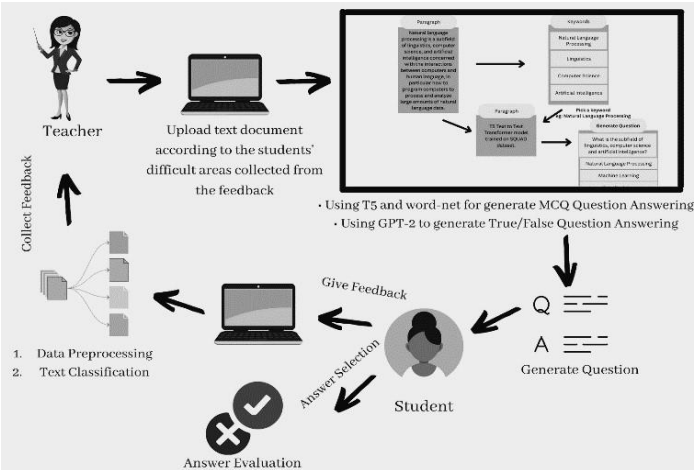

Figure 5: - System Overview Diagram Automated Personalized Question Generation and Answer Evaluation

For automated chat bot system for answering student questions can be developed with BERT transformer model which will take the notes to learn then model will take the question as input to answer the questions. BERT (Bidirectional Encoder Representations from Transformers) is a deep learning model in which every output element is linked to every input element and the weightings between them are dynamically determined depending on their connection. Historically, language models could only interpret text input sequentially from left to right or right to left, but not both. BERT is unique in that since it can read in both directions at the same time. This feature, made possible with the introduction of Transformers, is referred to as bi-directionality [18].

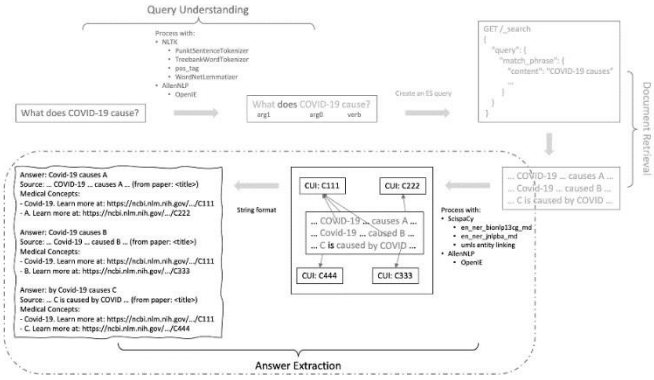

Figure 6: - NLTK model's way of answering the question.

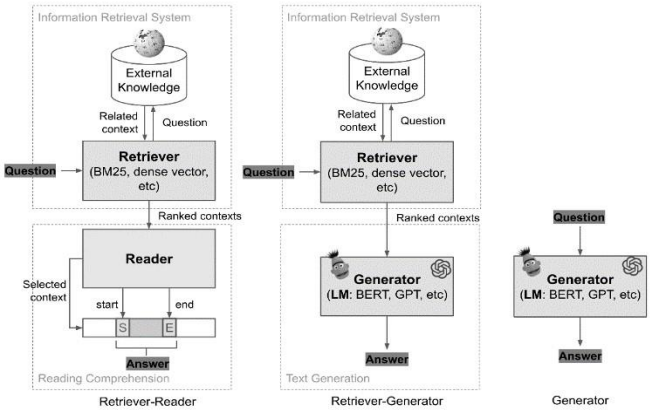

Figure 7: - BERT functional structure

5. Results

This system will be used as a web application for the online video classroom app that makes teaching and learning easy. After all the required setup and registration process, the students will be able to join the online classes. The system will start monitoring the students once the teacher enables the feature and it will be giving alert to the students who have been identified as being distracted. The results of the complete product were calculated separately under the respective components. Initially, the system that will detect that a student's eyes are closed for a few seconds. This system will alert when drowsiness is detected.

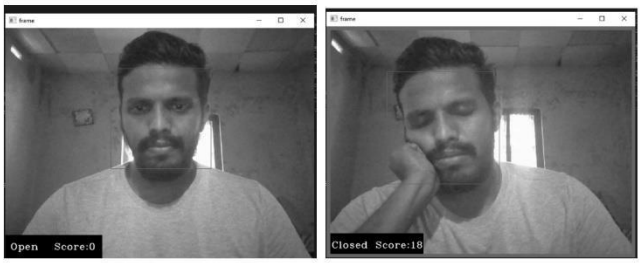

Figure 8: - Results Output of Drowsiness Detection and Alerting the Student

Moreover, system read the audio file which is having background noise and analyze the audio signal. Then it detects the noisy signal and remove it from the original audio signal. Finally get an audio without background noise. After that the system generate the transcript from the noise filtered audio file.

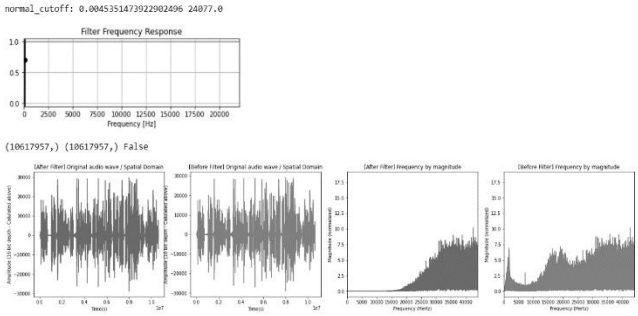

Figure 9: - Noise Removal

Also, after getting feedback from the students on how much they had understood the lesson, the system will be able to generate personalized questions from the notes that were uploaded to the system by the teacher for their practice. Inferences are shown after the findings in the table below.

Sentence No	No of sentences	No of correct questions	No of incorrect questions	No of question by human
1	1	2	0	2
2	1	2	0	2
3	2	3	1	2
4	1	2	0	2
5	1	2	0	2
6	1	2	0	2
7	2	1	1	3
8	1	2	0	2
9	1	2	0	3
10	1	3	0	3

Table 1: - Summary of Generated Questions

Total no of sentences = 12
Total no of correct questions = 21
Total no of incorrect questions = 2
Total no of questions by the human = 23
The above result shows that the system is working fairly correct with an accuracy of over 90% which can be further improved.

Furthermore, the students can also use the chat-bot to clear their doubts or get answers for the questions they need in a quick time.

Data used for Training	Approach	F1 Score
ClariQ	Full Dataset including "Impossible Answer"	11.34
ClariQ	Uses snippets of abstracts	25.22
ClariQ	Full Abstract without "Impossible Question" feature	62.23
ClariQ + SQuAD	Full Abstract without "Impossible Question" feature	74.57

Table 2: - F1 score for different approaches

As the table shows, the initial strategy resulted in a low F1 score. This outcome was enhanced by extracting and using abstract snippets from reference materials. However, the outcomes were still dismal at this stage. After extracting entire abstractions, the model significantly improved and began to exhibit a modest ability to answer questions. Results improved once again after integrating an extra QA dataset from SQuAD in training.

6. Conclusion

This paper discusses the technologies and development strategies of our system and, how it could help for the betterment of e-learning. The tactics of four different components were analyzed in order to produce a list of guidelines for future research in this field. Numerous research projects have been carried out in the topic 'e-learning'. They have mainly focused on provide more convenient platform for the teachers and students by reducing their time and effort spent on the tasks. The main objective of the system works by producing an online video classroom web application that monitors, manages, and enhances the online education system's efficiency. The sub objectives of the system are machine learning approach for students' attention of the class through video processing, automatic video transcription according to user selected language, automated personalized question generation and answer evaluation and a chat-bot that helps students to get answers to questions within the syllabus quickly and easily.

References

[1] Z. S. N. M. A. a. M. M. Mardi, "EEG-based drowsiness detection for safe driving using chaotic features and statistical tests," *Journal of Medical Signals And Sensors*, vol. 1, no. 2, p. 130, 2011.

[2] S. M. R. a. M. M. Noori, "Driving drowsiness detection using fusion of electroencephalography, electrooculography, and driving quality signals.," *J Med Signals Sens*, vol. 6, no. 1, p. 39, 2016.

[3] J. D. S. U. T. D. E. a. M. G. Krajewski, "Steering wheel behavior based estimation of fatigue," in *Driving Assessment Conference 2009*, 2009.

[4] C. B. a. C. Downey, "Automated generation of 'good enough' transcripts as a first step to transcription of audio-recorded data," *Sage Journals*, vol. 11, no. 2, 2018.

[5] M. H. Kaustubh Bhattacharyya, "Speech Background Noise Removal Using Different Linear Filtering Techniques," 2018.

[6] D. Y. S. R. I. Patel, "Speech Recognition Using HMM with MFCC-AN Analysis Using Frequency Spectral Decomposition Technique," *Signal & Image Processing : An International Journal(SIPIJ)*, vol. 1, p. 10, 2010.

[7] S. S. Nitin Washani, "Speech Recognition System," *International Journal of Computer Applications*, vol. 115, p. 8, 2015.

[8] G. a. M. K. S. Nijhawan, "Real Time Speaker Recognition System for Hindi Words," *I.J. Information Engineering and Electronic Business*, vol. 6, no. 2, p. 35, 2014.

[9] B. M. M. P. S. &. S. A. A. Das, "Automatic question generation and answer assessment: a survey," *Research and Practice in Technology Enhanced Learning*, p. 15, 2021.

[10] I. K. H. a. D. M. Feinerer, "Text mining infrastructure in R," *Journal of statistical software*, p. 54, 2008.

[11] A. S. A. S. B. N. A. &. N. S. J. Sreelakshmi, "A question answering and quiz generation chatbot for education," in *Grace Hopper Celebration India (GHCI) - IEEE.*, 2019.

[12] A. Sarvaiya, "An Intel Developer Program Documentation," jul 2018. [Online]. Available: https://software.intel.com/en-us/articles/using- natural language processing forsmart-questiongeneration.

[13] N. N. J. S. J. P. S. a. R. C. Bhartiya, "Artificial neural network based university chatbot system," in *IEEE Bombay Section Signature Conference (IBSSC)*, Mumbai, 2019.

[14] K. A. C. A. P. A. C. a. S. D. Shingte, "Chatbot Development for Educational Institute," 2021.

[15] T. Davies, "Knowledge Bases and Neural Network Synthesis," in *Pacific Rim International Conference on Artificial Intelligence*, 1991.

[16] G. Romao, "A Literature Review on chatbots in education: An intelligent chat agent," in *Universidade Católica de Moçambique*, 2021.

[17] M. H. F. K. H. A. a. C. C. Karam, "Noise removal in speech processing using spectral subtraction," *Journal of Signal and Information Processing*, vol. 5, p. 10, 2014.

[18] E. R. R. D. M. M. a. R. S. Fu, "BERT for Question Answering on," *SMU Data Science Review*, vol. 3, p. 15, 2020.

Author Profile

Soujanyaah Karunanithy
Corresponding author
BSc (Hons) in Information Technology (Reading)
SLIIT

Kishome Sivarajah
BSc (Hons) in Information Technology (Reading)
SLIIT

Viknatharshan Nadarajah
BSc (Hons) in Information Technology (Reading)
SLIIT

Sivashankar Sivaganeshamoorthy
BSc (Hons) in Information Technology (Reading)
SLIIT