**“AITor” - EDUCATION PLATFORM – A PERSONALIZED STUDENT PERFORMANCE ANALYZER AND RECOMMENDATION SYSTEM**

2022-017

Final Report

Liyanage M.L.A.P.– IT19120812

B.Sc. (Hons) in Information Technology specializing in Software Engineering

Department of Computer Science and Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

September 2022

**“AITor” - EDUCATION PLATFORM – A PERSONALIZED STUDENT PERFORMANCE ANALYZER AND RECOMMENDATION SYSTEM**

2022-017

Final Report

Liyanage M.L.A.P.– IT19120812

B.Sc. (Hons) in Information Technology specializing in Software Engineering

Department of Computer Science and Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

September 2022

# DECLARATION

We declare that this is our own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

|  |  |  |
| --- | --- | --- |
| Name | Student ID | Signature |
| Liyanage M.L.A.P. | IT19120812 | Shape, arrow  Description automatically generated |

The above candidates are carrying out research for the undergraduate Dissertation

under my supervision.

Signature of the supervisor: Date:

2/11/2022

…………………………………….

# ABSTRACT

With the transition to the “new normal” due to the COVID-19 pandemic, most of the human interactions were limited. With less physical interactivity, society is gradually adapting more online ways to fulfill its daily requirements. Among them, online delivery of education can be highlighted. Due to the virtuality and less interactiveness with the tutor and children, online education has many deprivations such as the inability to measure students’ engagement, productivity, feedback, and many more compared to in-class education. Refer to [1], [2], where it is well-stated that learning style, teaching competency, learning resources are positively contributing towards achieving high productivity on online education.

This case study will focus on identifying and analyzing the student’s performance on online education on an individual basis and provide personalized recommendations on study materials. Specifically, this will focus on identifying, predicting, and recommending the best learning strategies so that students can have the most suited learning style and teaching competency. The study will assign the student with a combination of pre-defined learning strategies such as “Video-based learning”, “Text-based learning”, “audio-based learning”. Apart from that, building up an “identification and recommendation” algorithm will cater to students with the best learning resources. Further to the above objectives, the study complies with an evaluation-based learning progress monitoring model to orchestrate the students to achieve high productivity with an intelligent feedback mechanism. To provide solutions to the above objectives, a web-based software solution known as “AITor” is proposed which will be acting as an interface between the student, the tutor, and the system. The application is included with a feature to capture the student in real-time and analyze his/her facial behavior when learning with the respective learning strategy.

In-conclusion, the proposed application is all-in-one learning management system, which provides a detaild view of the student engagement and manoeuvre students with proper guidance and identified industry trends.

CONTENT

[DECLARATION iii](#_Toc117241003)

[ABSTRACT iv](#_Toc117241004)

[LIST OF FIGUERS vii](#_Toc117241005)

[LIST OF TABLES viii](#_Toc117241006)

[LIST OF ABBRIVIATIONS ix](#_Toc117241007)

[1. INTRODUCTION 1](#_Toc117241008)

[1.1. Background and Litereature Survey 1](#_Toc117241009)

[1.1.1. Background review 1](#_Toc117241010)

[1.1.2. Literature review 5](#_Toc117241011)

[1.2. Research Gap 9](#_Toc117241012)

[2. Research Problem 12](#_Toc117241013)

[3. OBJECTIVES 14](#_Toc117241014)

[3.1. Main Objectives 14](#_Toc117241015)

[3.2. Specific Objectives 15](#_Toc117241016)

[3.2.1. Identification and prediction of best learning strategy 15](#_Toc117241017)

[4. METHODOLOGY 17](#_Toc117241018)

[4.1. Introduction and basic user flow 17](#_Toc117241019)

[**4.2.** **Dynamic weighted probability allocation model** 20](#_Toc117241020)

[4.3. Proposed Technologies 24](#_Toc117241021)

[4.4. Requirement gathering 26](#_Toc117241022)

[4.4.1. Data gathering and pre-processing 26](#_Toc117241023)

[4.4.2. What is proposed to achieve from the data set 27](#_Toc117241024)

[4.5. Feasibility study 27](#_Toc117241025)

[4.5.1. Implementation feasibility 27](#_Toc117241026)

[4.5.2. Technical feasibility 27](#_Toc117241027)

[4.6. Implementation 27](#_Toc117241028)

[4.7. Deployment 28](#_Toc117241029)

[5. DESCRIPTION OF PERSONAL AND FACILITIES 29](#_Toc117241030)

[5.1. Usage of cloud infrastructure 29](#_Toc117241031)

[6. Project Requirments 30](#_Toc117241032)

[6.1. Functional requirments 30](#_Toc117241033)

[6.2. Non-functional requirments 30](#_Toc117241034)

[7. BUDGET AND BUDGET JUSTIFICATION 31](#_Toc117241035)

[7.1. Cloud Cost 31](#_Toc117241036)

[7.2. Marketing cost 31](#_Toc117241037)

[8. Gantt Chart 32](#_Toc117241038)

[REFERENCE LIST 33](#_Toc117241039)

# LIST OF FIGUERS

[Figure 1.1.1 : Summary of responces of comfortability of online delivery of education 2](#_Toc117241042)

[Figure 1.1.2 - Variation of teaching methods used by tutors 3](#_Toc117241043)

[Figure 1.1.3 - Student's engagement in an online session 3](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2022RegCloud/2022-017-Students/1.%20Project%20Proposal/Report/IT19120812_2022-017_Proposal_Document.docx#_Toc117241044)

[Figure 1.1.4 - Student's engagement methods/learning strategies 4](#_Toc117241045)

[Figure 1.1.5 - Model accuracy test results 7](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2022RegCloud/2022-017-Students/1.%20Project%20Proposal/Report/IT19120812_2022-017_Proposal_Document.docx#_Toc117241046)

[Figure 1.1.6 - Accuracy comparision with proposed new model and existing models 7](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2022RegCloud/2022-017-Students/1.%20Project%20Proposal/Report/IT19120812_2022-017_Proposal_Document.docx#_Toc117241047)

[Figure 1.2.1 - Number of hours students are spending on the online education 12](#_Toc117241048)

[Figure 3.1.1 - responses received for the question asked related to the teacher’s guidance in online education 14](#_Toc117241049)

[Figure 4.1.1 - Pre-evaluation stage 18](#_Toc117241050)

[Figure 4.1.2 - Basic user flow 19](#_Toc117241051)

[Figure 4.2.0.1 - supportive materials uase in online education 21](#_Toc117241052)

[Figure 4.2.0.2 - Learning strategies 21](#_Toc117241053)

[Figure 4.2.0.3 - core hypothesis function 22](#_Toc117241054)

[Figure 4.3.1 - proposed way of identifying possibility weighting factors 25](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2022RegCloud/2022-017-Students/1.%20Project%20Proposal/Report/IT19120812_2022-017_Proposal_Document.docx#_Toc117241055)

[Figure 4.3.2 - Proposed posibility weightning factor algorithem 25](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2022RegCloud/2022-017-Students/1.%20Project%20Proposal/Report/IT19120812_2022-017_Proposal_Document.docx#_Toc117241056)

[Figure 4.7.1 - High level deployment method 28](#_Toc117241057)

[Figure 7.2.1 - Gantt Chart 32](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2022RegCloud/2022-017-Students/1.%20Project%20Proposal/Report/IT19120812_2022-017_Proposal_Document.docx#_Toc117241058)

# LIST OF TABLES

[Table 0.1 - table of abbriviations ix](#_Toc95481725)

[Table 1.1 - Reseatch gaps of the identified case studies 11](#_Toc95481726)

[Table 1.2 - Research gap in live applications 11](#_Toc95481727)

[Table 6.1 - Cloud budget justification 28](#_Toc95481728)

[Table 6.2 - Marketing budget justification 28](#_Toc95481729)

# LIST OF ABBRIVIATIONS

Table 0.1 - table of abbriviations

|  |  |
| --- | --- |
| Abbriviation | Description |
| **LMS** | Learning Management System |
| **SARS-COV-2** | severe acute respiratory syndrome coronavirus 2 |
| **OL** | Observational Learning |
| **SRL** | Self Regulated Learning |
| **ML** | Machine Learning |
| **DERN** | Deep Engagement Recognition Network |
| **LSTM** | Long Short Term Memory |
| **LRCN** | Long-term Recurrent Convolution Network |
| **PPWF** | Posibility Percentage Weight Factor |

# INTRODUCTION

## Background and Litereature Survey

### Background review

With the emergence and speading of Sars-Cov-2, most of the countries were taken actions to limit the human interractions fully or partially. With the enforcement of new rules and regulations, most of the day-to-day activities such as transportation, supply-chain and logistics, education were bounded. According to the UNESCO’s monitoring report [3], more than 160 countries worldwide enacted national shutdowns, affecting approximately 87 percent of the worlds student population. Due to this, with the aid of advancemet of technology, most of the schools, higher educational institutes and related parties have adopted to the online delivery of their course materials.

With the high rate of this adoption, a high demand for software solutions like Learning Management Syatems (LMS), E-learning portals (Microsoft SharePoint), document/content sharing applications were created. Most of the schools, higher educational institutes, universites have provided access to their own LMS. Furthermore, according to [4], video conferencing tools such as Zoom, Microsoft Teams, Google Meet have received a significant demand for their services due to delivery of live lectuers.

Nevertheless, with the immediate transition from face-to-face classroom to a virtual one, most of the students and educators founds it dificult to adjust. Considering the inability to measure and monitor the student’s actions in a virtual environment, most of the educators were unable to assess his/her student’s progress [5]. With regards to the student’s feedback, according to the [6], reduced focus, psychological issues, and management issues were identified.

Inorder to test some hypothesis, a closed survey was conducted. An online questionire was distributed among 57 school tutors, tutors of higher educational institutes to get their feedback about above issues identified. As shown in Figure 1.1, majority of the tutors are comfortable with delivering education online.

Chart, bar chart

Description automatically generated

Figure 1.1.1 : Summary of responces of comfortability of online delivery of education

According to [4], there are enough resources if someone is planning to transition from ordinary education methods to online education. Eventhough the availability of the resources are high, the adoption to those technologies can make extra effort in some scenarios.

However, paying attention to the percentages shown in the Figure 1.2, it is identifiable that most of the tutors are stick to less number of varied teaching methods in online education. Explaining further, these teaching methods are the tools thay use, the content type thay use to explain a theory concept (video, audio, text etc).

Focusing on the student’s engagement, according to Figure 1.3, majority of the students were moderately engaging with the session while second-majority was slightly engaging. Due to this, it is fair to say that, student’s performance and the productivity of the education can be alterd with the transition of online education.

Chart, bar chart

Description automatically generated

Figure 1.1.2 - Variation of teaching methods used by tutors

Chart, pie chart

Description automatically generated

Figure 1.1.3 - Student's engagement in an online session

Aiming the student’s issues stated in the [6], a separate closed survey was conducted among 48 university students. A question was enclosed in the survey to identify the most engaging activities students are might fond of. Intention of including this question is to identify the potential solutions to boost the focus on subject mattes, learning in the online education. Refer the Figure 1.4 for more details.

Chart, bar chart

Description automatically generated

Figure 1.1.4 - Student's engagement methods/learning strategies

According to the responses received, superiority of students were interested in video based learning where same university lecturer or another tutor explaining the theory based concepts. It is noticeable that, the second majority of the sample were interested in learning subject matters with text based resources like reference books, E-books etc. least number of students were attentive with the audio based learning concept. This includes podcasts and live lectueres with only audio.

Briefing the Figure 1.4, a hypothesis can build up stating different students have different learning strategies/engagement methods when it comes to online education. Since this is directly affecting towards the productivity of the education, it is worthy to add this as a key objective for the end solution.

Contrasting the background review, both the tutors/teachers and students have found it difficult to increase the engagement in online education. Form the tutor’s/teacher’s point of view, the engagement could be limiting due to the less variance of the teaching methods while from the student’s point of view it could be the high variance of the learning strategies/engagement methods.

### Literature review

As mentioned in the background review, identifying the best learning strategys/engagement methods can positively impact the productivity of online education. With the advancement of the machine learning, deep learning and cognitive sciences, new ways of human interactions with the computer have broaden up. With respect to the improvements of the above, sectors such as education, health Et al has gain momentous adjustments to its requirements. Transformation of classical teaching/learning system into an online based education resulted high usage of above-mentioned technologies in the education sector. With the help of the speed and the accuracy of different prediction and analytical algorithms available, we are able to identify peculiar criteria of student’s education individually and uniquely. Either combining several algorithms or tuning them, we can firmly monitor how the student is interacting with the course materials with the help of online means and make the suitable adjustments to cater new requirements in a feasible manner. There are several studies, that have been done with the objective of finding and fine-tuning best learning strategies in the online education. Here are some of the case studies done on the same research area.

Yang Tzu-Chi has done research on how observational learning and self-regulated learning strategies can affect the online learning performance of student [7]. The case study has focused on how observational learning (OL) and self-regulated learning (SRL) can link up with the online learning strategies. With the research context, it was identified that, the learning performance can be positively altered with proper identification of behavioral patterns of the student. [7]

Furthermore, the implemented system has dual proposed mechanisms to support both OL and SRL [7]. It includes, setting and externalizing the goals, planning the learning strategies and the time student is willing/planning to study a particular subject area. The study was conducted in a classroom, with 2 main assessment tests as pre and post tests. All the students were made aware about the procedure which includes weekly online tests, after each lecture was delivered. Before the weekly assessments, the students were made to face a classic-type examination and after the online tutoring, there was another post-test to attend. With the experimental results, it was highlighted that there is no clear relationship between OL and SRL. Furthermore, it has identified that, students with SRL strategies have high performance in learning.

Tao Huang et al, has done another case study on Fine-grained engagement recognition in online education [8]. This research is based on facial features recognition in account with the student’s live engagement to the course materials which helps the tutors to get timely feedback of their students. The case study itself has proposed a new machine learning (ML) model which is known as Deep Engagement Recognition Network (DERN) which combines temporal convolution, bidirectional LSTM and attention mechanism [8]. With their proposed solution, we can extract facial features including eye gaze direction, head posture, face movement, facial expressions and eye changes. they have used Openface[9] to extract the features from the face and predict the above mentioned engagement activities. However, finding these features from a live video footage makes it bit difficult to handle the amount of data that will be collected. To overcome the scenario, a temporal convolution process is being used which will chunk down the video and extract the similar frames to extract the features. This is being done by two sub temporal convolution sectors namely, shallow convolution and deep convolution. From the shallow convolution, it will directly process the raw video footage and tries to identify similar changes of different frames. Then the deep convolution will analyze the output of the shallow convolution to sharply identify and build high-order relationship between features. [8]. This case study has used DAISEE [10] data set and has used five models including InceptionNet Frame Level, InceptionNet video level, C3D training, C3D Fine Tuning and LRCN to perform the experiment of the engagement recognition on the dataset. In-order to verify the validity and the accuracy of the model, they have used five-fold cross validation for binary classification, three classification and four classifications. According to Figure 1.5, it has been identified that LRCN model with four classifications has achieved 57.9% accuracy.

Table

Description automatically generatedSource: [8]

Figure 1.1.5 - Model accuracy test results

Table

Description automatically generated

Figure 1.1.6 - Accuracy comparision with proposed new model and existing models

Source: [8]

Jim B.J.Huang et al has done a research on exploring the learning strategies by sequence clustering and analyzing their correlation with student’s engagement and learning outcome. [11]. This case study has followed and iterative process of collecting and analyzing data to fine-tune the identification of the strategies. They have used, pre-defied questions and interviews to collect the data. With the test results, they were able to identify the learning strategy by the logs of the student test results, identify the relationship between learning strategy and learning outcome. The research was done with 53 college students with the data of python course. the questionnaire developed by the Elmaadaway [12] has made with 3 sectors covering behavioral, cognitive and emotional engagement.

Fidelia O. and Et al have done research [13] on how finding a relationship between students engagement and hei/her performance. Thay have primareliy focused on Total time spent in MindTap (TimeOnTask), Number of logins, Percentage of Activities Accessed featues when building up the relationship. They have used both supervised learning (Random forests) and unsupervised learning (clustering) models to fine-tune the final output.

In contrast, Based on the research done in the above sector, it is identifiable that, in order to achieve a high accuracy of identifying and predicting the best learning strategy of the student, usage of combination of few methods is a necessity. With the behavioral analysis, assignment grading review and background check, the proposed solution will achieve a high accuracy rate.

## Research Gap

Keeping the objective as finding the best learning strategy, reserches have done many case studies on it. Referring to the methodologies and the final findings included from [7]-[11], and [13] most of the case studies either have followed a full technical methodologies such as

* behavioral analysis
* facial analysis

or full old-schooled methodologies such as;

* questioners
* face-to-face interviews
* number of hours worked on a task
* make decisions depending on the assessment marks, the students have received.

Another noticeable feature that can be highlighted from the above studies is, none of the case studies have mainly focused on building up a relationship between the student’s engagement and the learning strategies. Most of the researchers have tried to build up a relationship between engagement and the final productivity of the students. In the [8] (will be known as A), researchers have done implemented a software solution to identify student engagement via a deep facial analysis. According to the [11] (will be known as B),

* Reading the syllabus
* Read before class
* Reading after class
* Past exam practice

Have been identified as learning strategies. But yet those identifiables have limited territory when it comes to the present content types, students are using to learn.

According to the survey done (Figure 1.4), it can be identified that most of the students have individual ways of learning the subject matter. With the previously done researchers, they have not linked the learning strategies with engagement increasing and productivity.

The below table contains a detailed view of how each case study ([8],[11]) has addressed the engagement analysis and their attributable outputs.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Case study | Does it has engagement analysis mechanism | Does it has a relationship between engagement analysis and learning strategies? | Is there any categorization of students according to their learning strategies? | Main output |
| A | yes | no | no | Students will be categorized as either engaging or not-engaging by perfoming real-time facial analysis. |
| B | yes | no | yes | Categorization of students according to their learning strategies and engagement level. |
| AITor | yes | yes | yes | Categorization of students according to their learning strategies and engagement level. |

Table 1.1 - Reseatch gaps of the identified case studies

Besides the above case studies, considering the real-world applications like Moodle,

Blackboard and distance learning platforms like Udemy, Coursera also do not have any feature to individually extract the best-suited learning strategy of the student. They have designed to provide the content in a manner of the same method to the entire user group. The below table contains a conclusion of how the above-mentioned real-world applications have implemented the provision of learning strategies to their users.

Table 1.2 - Research gap in live applications

|  |  |  |
| --- | --- | --- |
| Platform | Identifiable learning strategies | Are the learning strategies personalized? |
| Udemy | Video, Text ,Live engagement activities like live-coding | No |
| Coursera | Video, Text, Live engagement activities like live-coding | No |
| AITor | Video, Audio, Text  Live engagement activities | Yes |

# Research Problem

Depending on the type of course and the program type, whether it's full-time or part-time, students are required to spend a considerable number of hours on the course content. With online education, this number may have been increased. From the closed survey done, it can be identified that most of the students are spending 10+ hours in their course content. Refer to Figure 1.7 for more information.

Chart, bar chart

Description automatically generated

Figure 1.2.1 - Number of hours students are spending on the online education

Furthermore, according to [14], students’ behavior and emotional engagement are highly effective to the final productivity of the education. In the context of [1], learning resources and learning style/strategy is a direct influencing factor for a highly effective online education.

Further to the results acquired by the survey, Figure 1.4 includes the different ways of learning strategies students are using to gather the required knowledge. Although predicting the best-suited learning strategy of each student is somewhat qualitative and highly volatile, this case study will aim at predicting it by combining user feedback analysis and rapid assessment evaluation methodology.

The case study will provide a high-level overview of how knowledge-inspired computing can be used in the educational sector in order to enhance, predict and analyze the students learning strategies in online education. By usage of pattern recognition and unsupervised learning algorithms, the AITor application is loaded with a technology-enhanced fine-grained learning strategy prediction feature.

# OBJECTIVES

## Main Objectives

The main objective of this case study is to provide a high engaging technology-enhanced online education experience to both the students and tutors where students get the ability to experience more personalized education and tutors to have an upfront detailed analysis of his /her students. Given a fact that most of the educational institutes are now delivering their course content online and according to the analysis done in the previous sections of this document, students have find it difficult to grasp the entire knowledge from the content delivered via the tutor. Not only that according to Figure 2.1, but the majority of the students also find the teacher’s guidance as modetatly healpfull when it comes to online education.

Chart, bar chart

Description automatically generated

Figure 3.1.1 - responses received for the question asked related to the teacher’s guidance in online education

As per the above explanations, from the student's perspective, the main objective of this study is to provide a better platform that is personalized to each individual and has a suitable recommendation of course content mechanism.

Furthermore, from the tutor’s perspective in itself, the main objective is to provide a platform that provides a detailed analysis of students.

Concerning the above main objective, the case study is being chunked down into further specific objectives. They can be listed as follows.

## Specific Objectives

### Identification and prediction of best learning strategy

This objective is to scheme out the mechanism of how the application is going to identify, predict and analyze the suited learning strategy/s of the student. Authors have to research and decide the available and most used learning strategies from the students. This includes building up and testing up the hypothesis of how to find the optimal learning strategy if each student. In most cases, one student may have an array of learning strategies (video, audio, text). This objective should address those requirements as well.

Furthermore, this objective includes the way of implementing the above features in the AITor web application. Authors will have to consider best practices and best UI/UX principles when implementing those features.

Apart from that, to address privacy, proper consent should be received from the user. This objective will address the proper provision of guidelines and methodologies that are up with current industry norms.

Those all requirements should be achieved through the web application and, to provide seamless implementation the features will be implemented by following agile methodologies. This will include incremental implementation with time-bounded deadlines.

# METHODOLOGY

## Introduction and basic user flow

The case study is proposing a full scalable web application that has the capability of;

* Analysing, identifying and predicting the best learning strategy/s of a student
* Recommending course related materials according to the identified learning strategies
* Asses the student with the standard assements provided by the tutor and analyse them.
* Identify the potential skills of the student related to computer science.
* Provide detailed analysis of each student as well as entire batch of the student to the tutor or to the institute

This section will mainly focus on how to implement the extraction of best learning strategy of the student via the web application.

With the knowledge gain from the literarture study, it was able to identify that, most of the case studies or applied researches have only considered implementing featuers to either extract the engagement percentage or the learning strategies. For engagement analysis, most of the studies have used facial emotional analysis and for learning stragey identification, majority of the studies have collected data via online questionires, past assessment grades and surveys. However it is clear that there should be a significant relationship between the engagement and the learning strategy.

The proposed feature includes combinantion of engagement analysis, questionire review and assessment review. The model will evaluate each input and based on statistical proof, the ideal learning strategy/s will generate. This actions will cope with generating highly accurate results compared to the studies done in the above.

Furthermore, the proposed model is consiste of two major sub-sections namely

1. Pre-evaluation stage
2. Post-evaluation stage

Figure 3.1 includes the high-level overview of how inputs of this two sub-section will implement and interact with each in the proposed web application.

Diagram

Description automatically generated

Figure 4.1.1 - Pre-evaluation stage

The post-evaluation stage is interacting with the marks get by the students in their day-to-day assessments. If the marks are getting low by the time, it will re-iterate the above process and will fine-tune the results get by the pre-evaluation stage.

Features and responsibilities of each sub-modules can be brief as follows.

1. Pre-evaluation stage

this feature is fully web based dynamic interface that the student will directly interacting with. If the user is new to the system, the user will navigate through a pre-defined path in order to get the basic information and surey data. A well defined consent will take in this stage, since the data gatherd in the basic information and the survey has some privacy concers. After a successful provision of above data and information, the student will be undergo an pre-defined set of proceduers. As the first step, the student will be taught a set of theory concepts upon the selection of the theory via the identified three means of learning strategies (video, audio, text). Then an immediate assesment will be provided to the student to assess the competency level of the student related to the content covered in the above three learning strategies. Upon assessing the assessment, the dynamic engagement analysis and learning strategy identification model will evaluate the student and predect the suitable weights/percentages of the learning strategies that the particular student should be labeled. Figure 3.4 will brief how the user flow will takes place in the above scenario.

Diagram

Description automatically generated

Figure 4.1.2 - Basic user flow

**Dynamic weighted probability allocation model**

To identify and assign necessary possibility elements to each of the aforementioned learning strategies, a novel Dynamic Weighted Probability Allocation (DWPA) method has been proposed in this case study. The new DWPA algorithm determines the necessary weights for the weighted function using logistic regression and the straightforward weighted arithmetic mean. The newly introduced algorithm will take into account student's health characteristics, such as hearing and vision disorders, as well as other focus-related challenges, while creating the necessary weight elements. Additionally, it was primarily concerned with the kind of device the learner is using to access the educational materials. The above key characteristics have been identified among the features mentioned in the section B with the help of "hyper-parameter tuning" concept. Each student receives a unique weight factor based on their health and other relevant factors.

**Data Collection:** Since the introduced DWPA algorithm is based on logistic regression, sample data has been collected from 388 students via an online survey conducted through Google forms. Majority of the sample were university undergraduates (95 percent) and the rest were school children from age 15 to 19. From the sample, 99 percent of the students are fully engaging in distance learning while the rest are engaging hybrid education.

Before dive into the creation of the hypothesis function, hyper-parameter tuning process was done in order to identify the contributing parameters. In the data collection phase, the audience was asked several questions to clearly identify the factors they are much concerned about in the online learning.

Chart, bubble chart

Description automatically generated

Figure 4.2.0.1 - supportive materials uase in online education

According to the Figure 4.2.1, it has been identified that the entire audience is using supportive materials in the online education.

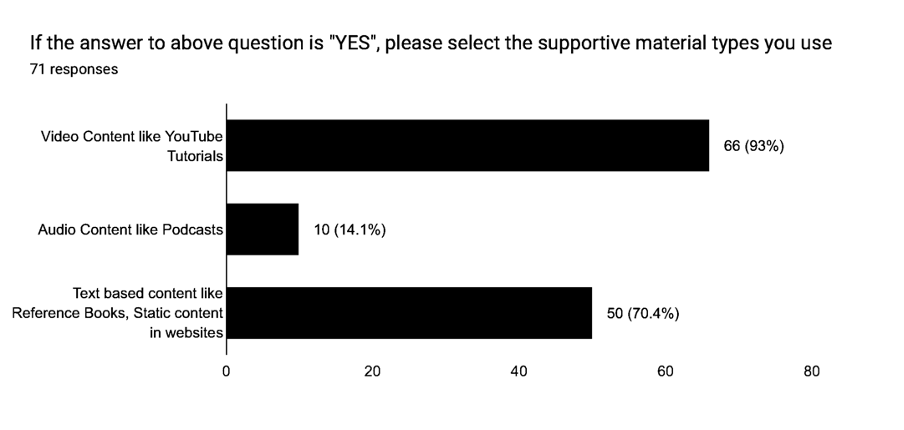


Figure 4.2.0.2 - Learning strategies

Furthermore, 3 main learning strategies has been identified as Audio, Video and Text. it can be concluded that, all the supportive materials that the students are using is falls under these three main strategies. Therefore, the case study has mainly focused on categorizing and optimising each students learning curve in-align with the above three main strategies.

Additional to the above, it was identified that, the compatibility of adoption to the online learning of students have been limited to few of the other parameters as well. According to the survey results, and the information gathered form the literature survey, below factors were identified as contributing factors to have a smooth online learning experience.

* Age
* Gender
* Specific health conditions related to learning
* Device type – D

From the survey feed-backs, it was identified that considerable number of students have Hearing(H), Vision(V), Focus related health issues(F). Therefore, specific health conditions have been further divided in to above sub-sections.

After identification of the contributing factors, with the help of "hyper-parameter optimization" technique, the above factors have been narrowed down to have a fine-grained hypothesis function. The hyper-parameter tuning phase concluded that, the aforementioned health conditions and the device type has a higher impact on the final probability factor.

Therefore, the newly introduced DWPA algorithm is based on linear polynomial function with these 4 distinct parameters(H, V, F, D).

Below is the primary arithmetic equation which is used to generate the probability weight of a given learning strategy. It is also the core hypothesis function which is used to calculate the probability weights of a given learning strategy.



Figure 4.2.0.3 - core hypothesis function

Here the alpha (a) value represents the the 3 learning strategies namely video, audio and text.

Shape

Description automatically generated with medium confidence

Figure 4.2.4 - Alpha values

But, given the fact that the P(a) value should a probability value and it should lie between 0 and 1, the aforementioned equation is being converted into a **sigmoid function**. This is where the principals of logistic regression is used in the DWPA algorithm. After the transformation of the above hypothesis function, below is the simplified version of the final equation.



Figure 4.2.5 - DWPA function

For more clarity, the plot for the newly introduced function will look like below.

Diagram

Description automatically generated with medium confidence

Figure 4.2.6 - function curve

Here, in order to find the best values for theta (f) **gradient descendent** method has been used. With the help of multivariate gradient descendent, the function will converge into the best possible theta (f) values. Since there is a requirement to fine-tune the weight factors (theta (f)) values of the student, the proposed algorithm will re-run the workflow with the updated values for H, V, F and D.

# TESTING AND IMPLEMENTATION

## Implementation

The implementation of the machine learning model is done using Python Django library. The traing data set is histed in awl S3 bucket and will be used to re-train the model or evaluate the accuracy of the model.

Text

Description automatically generated

## Deployment

The proposed application will be fully cloud native and will use Amazon Web Servises(AWS) as the go-to cloud provider. The application will use managed services of AWS to achive high availability and regulatory. Since the application is planned to adopt with DevOps a proper CI/CD imlmenetation is propsed to implement. All the components will be containerized with the help of Docker technologies.

Diagram

Description automatically generated

Figure 4.7.1 - High level deployment method

# RESULTS AND DISCUSSION

## Workflow accuracy and justifications

Focusing of the accuracy of the workflow, it has been optimised with the help of multivariate gradient decedent algorithm. The motivation behind the workflow is to fine-tune the assigned learning strategy probability weights each after the student is assigned with new H, V, F, D values. Therefore, the model will explore unique ways of reaching the global optima of the given data set more frequently resulting the perfect adaptation to the data set of the model Furthermore, since the proposed core equations is based on weighted average principals, the error impact on a given scenario will be minimized.

## Results

After successful implementation of the model, the application was able to predict the ideal learning strategy of each and every student with an post-signed accuracy level. This post-signed accuracy level will be used to re-evaluate and re-assign the possibility factiors to the students. Not to mention, the overall accuracy of the model has reached 86.67% and this will be further tuned with the re-evaluate phases.

## Reserch findings

in paralle to the literature survey done, apart from the variables that has been included in the DWPA algorithem, there are many factors, that can be consided to improve the productivity of distance learning. Further, this case study can be focused on particular sector to specialize and provide facilities with. As an example, this case study can be focued on personalizing and providing distance education to children with ADHD etc.

# CONCLUSION

Students engagement in the education is a major contributing factor towards the best performance no matter whether is distance learning or in-class learning. Yet in the distance learning, given the fact that the tutor/teacher do not have full closure with his/her students, boosting the engagement of the students is a must. Withing this case study, a novel algorithm called Dynamic Weighted Probability Allocation (DWPA) is being introduced. The intention of this algorithm is to fetch the students with personalized learning materials in order to increase the productivity of education. The output of aforementioned function will give fine-tuned probability weights for each learning strategy. Apart from that, with the help of gradient descendent algorithm, the values has been fine-tuned.

# REFERENCE LIST

[1] P. D. Barua, X. Zhou, R. Gururajan, and K. C. Chan, “Determination of Factors Influencing Student Engagement Using a Learning Management System in a Tertiary Setting,” in *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Dec. 2018, pp. 604–609. doi: 10.1109/WI.2018.00-30.

[2] A. Dwivedi, P. Dwivedi, S. Bobek, and S. Sternad Zabukovšek, “Factors affecting students’ engagement with online content in blended learning,” *Kybernetes*, vol. 48, no. 7, pp. 1500–1515, Jan. 2019, doi: 10.1108/K-10-2018-0559.

[3] “Impact Of Sars-Cov-2 And Its Reverberation In Global Higher Education And Mental Health,” *Psychiatry Research*, vol. 288, p. 112977, Jun. 2020, doi: 10.1016/j.psychres.2020.112977.

[4] C. K. Massner, *The Use of Videoconferencing in Higher Education*. IntechOpen, 2021. doi: 10.5772/intechopen.99308.

[5] M. Papadopoulou, P. Argyri, and Z. Smyrnaiou, “Online Education during COVID19 Papadopoulou Argyri Smyrnaiou 2021 ISTES,” 2021.

[6] M. Maqableh and M. Alia, “Evaluation online learning of undergraduate students under lockdown amidst COVID-19 Pandemic: The online learning experience and students’ satisfaction,” *Children and Youth Services Review*, vol. 128, p. 106160, Sep. 2021, doi: 10.1016/j.childyouth.2021.106160.

[7] Y. Tzu-Chi, “Impacts of Observational Learning and Self-regulated Learning Mechanisms on Online Learning Performance: A Case Study on High School Mathematics Course,” in *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, Tartu, Estonia, Jul. 2020, pp. 194–197. doi: 10.1109/ICALT49669.2020.00063.

[8] T. Huang, Y. Mei, H. Zhang, S. Liu, and H. Yang, “Fine-grained Engagement Recognition in Online Learning Environment,” in *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, Jul. 2019, pp. 338–341. doi: 10.1109/ICEIEC.2019.8784559.

[9] T. Baltrusaitis, A. Zadeh, Y. C. Lim, and L.-P. Morency, “OpenFace 2.0: Facial Behavior Analysis Toolkit,” in *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, May 2018, pp. 59–66. doi: 10.1109/FG.2018.00019.

[10] A. Gupta, A. D’Cunha, K. Awasthi, and V. Balasubramanian, “DAiSEE: Towards User Engagement Recognition in the Wild,” *arXiv:1609.01885 [cs]*, Apr. 2018, Accessed: Jan. 03, 2022. [Online]. Available: http://arxiv.org/abs/1609.01885

[11] J. B. J. Huang, A. Y. Q. Huang, O. H. T. Lu, and S. J. H. Yang, “Exploring Learning Strategies by Sequence Clustering and Analysing their Correlation with Student’s Engagement and Learning Outcome,” in *2021 International Conference on Advanced Learning Technologies (ICALT)*, Jul. 2021, pp. 360–362. doi: 10.1109/ICALT52272.2021.00115.

[12] M. Elmaadaway, “The effects of a flipped classroom approach on class engagement and skill performance in a Blackboard course: Effects of the flipped classroom approach,” *British Journal of Educational Technology*, vol. 49, Mar. 2017, doi: 10.1111/bjet.12553.

[13] F. Orji and J. Vassileva, “Using Machine Learning to Explore the Relation Between Student Engagement and Student Performance,” in *2020 24th International Conference Information Visualisation (IV)*, Sep. 2020, pp. 480–485. doi: 10.1109/IV51561.2020.00083.

[14] L. Yuwen and G. Yanping, “Research on the Promotion of Students’ Online Learning Engagement Based on Association Cluster Analysis,” in *2021 9th International Conference on Information and Education Technology (ICIET)*, Mar. 2021, pp. 247–251. doi: 10.1109/ICIET51873.2021.9419605.

[15] T. Huang, Y. Mei, H. Zhang, S. Liu, and H. Yang, “Fine-grained Engagement Recognition in Online Learning Environment,” in *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, Jul. 2019, pp. 338–341. doi: 10.1109/ICEIEC.2019.8784559.

[16] M. Elmaadaway, “The effects of a flipped classroom approach on class engagement and skill performance in a Blackboard course: Effects of the flipped classroom approach,” *British Journal of Educational Technology*, vol. 49, Mar. 2017, doi: 10.1111/bjet.12553.