Assessment of Academic Performance with the e-mental health interventions in virtual learning environment using Machine learning Techniques: A Hybrid approach

A Sheik Abdullah[1], RM Abirami[2], A Gitwina[3], C Varthana[4]
[1,2,3,4]Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamil Nadu.
[1]asait@tce.edu
[2] abiramiramanathan99@gmail.com
[3] navisgitwina@gmail.com
[4] varthanachandrasekar23@gmail.com

Abstract:

Background: The act of virtual learning is defined through learning and practicing in an environment using digital/electronic content for self-paced through online teaching and mentoring. It explicitly deals with the interaction in an asynchronous mode of learning. The quality of teaching-learning depends on the utilization of digital technologies with the advancement in educational technology. There is a need and evaluation for the assessment and estimation of the impact of e-mental health interventions with the students learning through the virtual learning environment.

Purpose/Hypothesis: This research evaluates the psychotherapeutic support for the students to overcome the psychological distress during this COVID-19 pandemic by using machine learning techniques. This mechanism evaluates the efficacy of the academic performance made by the students during the pandemic situation. This analysis involves a hybrid approach for the assessment in machine learning using a genetic algorithm with an artificial neural network upon statistical evaluation. The psychological factors are determined with a keen focus on behaviourism, cognitivism, and social constructivism. The metrics have been evaluated based on digital technologies (ICT) in remote access, individual learning process, flexible learning, cost-effectiveness, time complexity and scalability.

Design/Method: The design process involves the 775 student responses with 27 attributes with differentiation of labels corresponding to behaviourism, cognitivism, and social constructivism. The preprocessed data is fed to genetic algorithm with processing parameters focusing crossover and mutation probability and then classified using artificial neural network. The estimation of academic performance is made using the techniques followed in virtual learning environment such as:

1. Online quiz (Quizizz platform) – Individual assessment
2. Flipped classroom activity - Individual assessment
3. MOOCs online courses – Individual assessment
4. Prototype design – Team activity
5. Research proposal – Team activity

From the assessment process the each of the student performance is evaluated with regard to the course outcome of individual student in the learning environment. The variation has also been observed with the applicability of ALS and traditional practice methods.

Results: The hybrid approach found to be good in the assessment and evaluation of academic performance and health interventions in terms of accuracy (88.18%), precision (94.69%), recall (92.24%), RMS error (0.202) and correlation (0.844) respectively. The statistical analysis and evaluation have been made using Fisher's F-Statistical test, and the P-value is found significantly to be $P<0.001$.
From the experiments, the factors that contribute towards web-based learning, blended learning, and online learning has been differentiated with the psychotherapeutic factors. A total of 775 samples have been used for analysis with the applicability of ICT tools and the pedagogical practices for the course. The factors contributing towards Behaviorism with a focus on interaction and response towards the learning environment plays a significant role in varying the academic performance of the student of about 20% in total learning rate varied significantly. Step-by-step analysis in virtual learning provides a good initiative for the student's community to have a variation in the learning process. Virtual learning is one of the good practices if the ICT in education, process and its principles adhere more efficiently.

Keywords: Virtual learning, e-mental health, Genetic Algorithm, Artificial Neural Network,

1. Introduction

Learning is a process that encourages acquiring competencies, knowledge, and skills to develop future

opportunities. Thus learning is considered as the fundamental pillars of society changes. Virtual learning is the act of learning via digital platforms. A Virtual Learning Environment is a platform to students for delivering learning materials through the web. These environments include assessment, student tracking, and collaboration and communication tools (Lekha and Megha 2020). Virtual learning is one of the good practices if the ICT in education, process and its principles adhere more efficiently (Nikkie 2016).

E-mental health refers to mental healthcare practices. E-mental health involves leveraging the internet and related technologies such as Smartphone apps, web sites, and social media to deliver mental health services. E-mental health plays an important role to make an effort to deal with this societal challenge (quafae et al., 2019).

The psychological factors of students are determined with a keen focus on behaviorism, cognitivism, and social constructivism. This analysis involves a hybrid approach for the assessment in machine learning using a genetic algorithm with an artificial neural network upon statistical evaluation. Meanwhile, this research work focus on the development of a hybrid approach for the determination of e-Mental health interventions during virtual learning environment. It also evaluates the factors that contribute towards the interventions and health disorders that formulates as a basic criteria for the health issues (Tai and Ting 2019).

2. Literature Survey

The work by the authors (Alberto et al., 2020) proposed Artificial Intelligence methodologies such as tree-based model and common classification techniques and results were analysed accurately in parameters such as Decision trees(0.705), Random Forest(0.755), Extreme Gradient Boosting(0.765), Multilayer(0.782) and Simple Perceptron, Hopfield and convolution learning networks, logistic-hidden neural network(0.783).

The work by the authors (FanYang and Frederick 2018) came up with three steps such as Student Attribute Matrix (SAM) for quantifying attributes, estimation based on classification Back Propagation Neural Network (BP-NN) , and based on BP-NN student progress indicators and attribute casual relationships were proposed for analysing the mental health in which results were perfect upon estimation(49.72) and more accurate on other parameters.

The incorporation of machine learning techniques provided an diverse change in feature selection and classification process. The work by the authors (Juan et al., 2019)applied the techniques ML(Machine Learning),CF (Collaborative Filtering) , RS (Recommendation System),ANN(Artificial Neural Network) to predict the student behaviour according to e-learning environment The goal is to denotes the knowledge as group of grades and able to identify the similarity among students.

The work by the authors (Jae-Eun Russell et al., 2020) suggested the implementation of Elements of Success (EOS) in difficult courses for easy grading. It predicts details about students' performance ,activity and involvement in each category like study skills ,EOS provides an opportunity to improve, estimated Grade and Logistic Regression is used to find whether Eos helps in predicting the estimated grade or not.

The work by the authors (Lokanath Mishra et al., 2020) proposed methodologies such as Quantitative and Qualitative to analyse the impression of students and teachers and results found that 6% uses YouTube and face book platform, 18% found online platforms, 87% teachers found telephonic online platform for teaching. Similar work has been conducted by a methodology using three components like Pre-assessment, PBL session, Post - assessment. After collecting feedbacks from the educators in different learning environment with respect to the components quantitative analysis is done to track the activity of educators (Akhila Joshi et al., 2020)

The work by the authors (Rajhans et al., 2020) designed a cross sectional survey with validated questionnaire to find deviations in student learning activities on e-learning mode and usual mode, Survey from 2018 to 2020 results in 93.58% students have preferred e-learning mode. Similarly the authors (Bolinskia et al., 2020) proposed methods such as identification, screening, eligibility criteria included in Meta analysis. According to Academic performance only a small amount of deviation is analysed in students as a result.

This research models out the factors that contributing towards behaviorism, cognitivism, and social constructivism that formulates towards e-mental health interventions. The approach has been designed in a hybrid way of learning for the assessment and evaluation of academic performance. Algorithms focusing on machine learning techniques have been deployed upon statistical evaluation.

3. PROPOSED METHODOLOGY

We obtained the students dataset from Jordanian University repository and used python for coding purpose. The input dataset is done pre-processing by removing the redundant values and replacing the missing values. The methodology involves three steps Clustering, Feature selection and classification. Clustering algorithm proposed is K-Means algorithm where the data are grouped into three different clusters A, B, C. The feature selection algorithm we used is Genetic Algorithm (Parkavi et al., 2018). The reduced features from the algorithm is given as input to the Feed Forward Neural Network classifier from which the superlative precisions are found and used in the evaluation of efficacy of the academic performance made by students during COVID-19. The following Figure 1 depicts

the proposed methodological workflow. Once the classified attributes are evaluated then the statistical evaluation has been made using factor analysis (Parkavi et al., 2017). Finally the interventions that correspond to learning are identified with valid proofs.

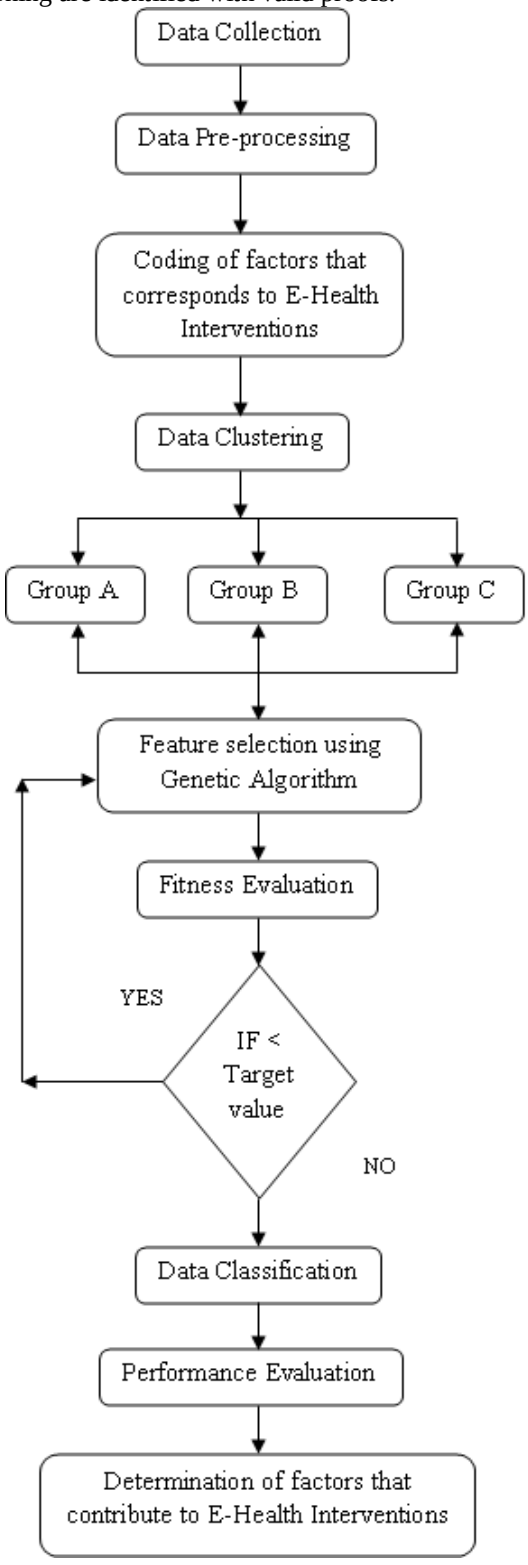

Fig. 1 Proposed Methodology

A. K-Means Algorithm Clustering

Clustering methods are used to identify groups of similar objects in a multivariate data set collected from various fields. Among the different types of clustering, K-Means clustering is preferred for getting the better results. It is an iterative algorithm which tries to partition objects into clusters in which each object belongs to the cluster with the nearest mean. This method produces exactly different clusters of greatest possible distinction (Larson et al., 2020).

B. Genetic Algorithm Feature Selection

Genetic Algorithm is generally based on natural selection and natural genetics is used for solving constrained and unconstrained optimization problems. It generally operates well on datasets of people to produce better and better approximations and also has the capability of delivering results at a high speed. The process includes creation of datasets, calculation of fitness functions, selection, crossover and mutation.

Generally Genetic algorithm is based on technique of Biological evolution .It is used where we need to understand the flexible process of natural system. Effective and efficient optimization can be achieved using Genetic algorithm .So it is used widely for business field, engineering field, scientific field and many other applications in machine learning. On the other hand it is also included in AI for getting copy of a human thinking, It is also a subset of Evolutionary Computation which is consider as the subfield of AI. Genetic algorithm is also used to overcome search problems, because easy to implement and also obtain better solution in search problems, It also find best solution for candidate sets that are broad and have many ideal points (Levine 2016).

C. Feed Forward Neural Network

Feed forward neural networks are used for classification and regression, as well as for pattern encoding. Their decision flow is unidirectional, advancing from the input to the output in successive layers, without cycles or loops. The feed forward model is the simplest form of neural network as information is only processed in one direction. Neural networks lead the way to the next step of Artificial Intelligence. It works as the same as human beings' neural networks. And it allows the system learning from the analysed data. Neural networks have a distinct capability to extract complex data and to find patterns and detect trends which is very complicated for the brains of humans and computers. Feed forward neural networks are the simplest form of neural network and do not form a cycle (Duarte et al., 2020). Here the Perceptions are set out and the input is in the input layer and output is generated in the output layer. Perceptron of every layer is linked with each unit in the next layer.

4. Experimental Results and Discussion

The data has been collected from https://data.mendeley.com/datasets/thnzm3yk23/1 with the analysis of e-learning and its risk factors that contribute towards virtual learning environment at Jordan University. Datasets are taken into consideration with initially 775 numbers of records. The first step of processing done is by removing the 15 missing datasets and the records obtained after it were 760. Duplicates were segregated among the remaining and 734 data were taken for the next process after the deletion of 26 duplicate datasets. Grouping the data based on their similarity into three different clusters A, B, C. Cluster A has 77 records, cluster B has 269 and cluster C has 388 records.

Table 1. Dataset and Processing Values

Dataset	Values	Total values after Processing
Original	775	775
Missing Values	15	760
Duplicate Values	26	734
K- Means Cluster	A	77
	B	269
	C	388

After Genetic Algorithm Feature Selection, the number of features in the subset is 19. The Percentile obtained is 0.55 and the Validation Accuracy is 0.80. The evaluation is made accordingly with the continuous values to that of the validation data. It has been estimated that the values are found to be significant. (Sheik Abdullah et al., 2018)

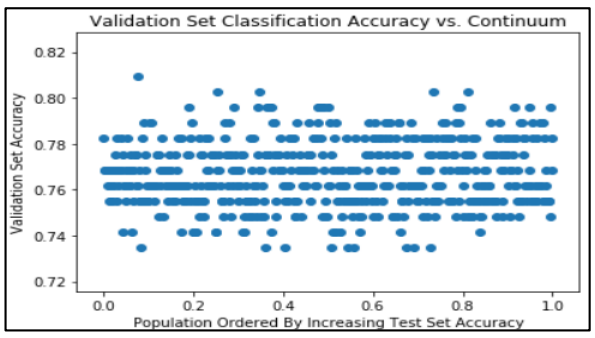

Fig. 2 Validation set classification accuracy vs continuum

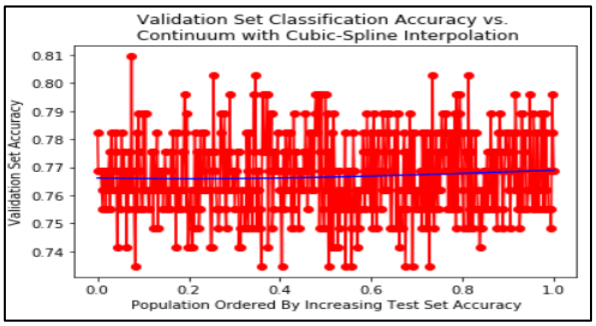

Fig .3 Validation Set Classification Accuracy vs. \n Continuum with cubic-Spline Interpolation

The accuracy obtained from the feed forward neural network is 88.18%.The Confusion Matrix obtained is depicted in Table 2. The graph in Figure 3 depicts the accuracy level with the validation data. The interpretation has been made using cubic-spline method of estimation. The estimation signifies that the values are predominant and found to be perfect with the observed data and its contributing factors.

Table 2.Confusion matrix

	C	A	B
C	107	0	6
A	0	17	5
C	9	9	70

The feedforward neural network is one of the biologically learned data classification algorithm. It includes a number of neuron-like units with an organized set of layers. All the data unites in the networks is found to be connected hence it formulates the name of feedforward network model.

Here we have used an two-layered network model which includes sigmoid hidden neuron-like units and an linear output frame. The entire model is trained using LMBP algorithm. The feed forward neural network view obtained as depicted in Figure 3 is developed with 12 intermediate layers for the signified levels A, B and C.

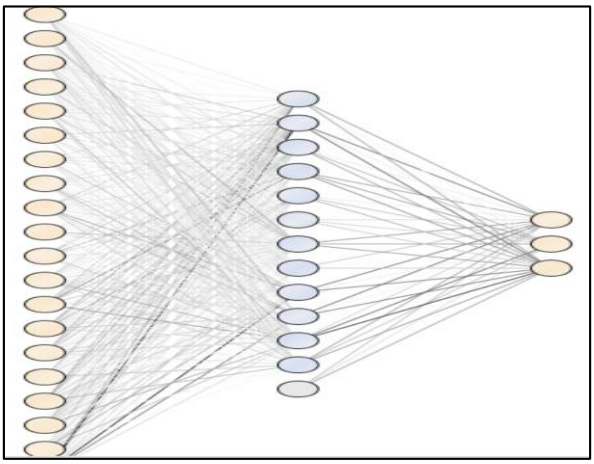

Fig.4 Feed forward neural network

5. Statistical Evaluation using Factor Analysis

In general, the dataset is said to have many forms of contributing factors which explains about the problem defined. During the stages of exploratory data analysis the realm of factor analysis can be made accordingly with the following separations:

- Factoring principal axis
- Maximum likelihood estimation

These forms of factor analysis are used with the determination of eigen values for each factor. This is then used to estimate the number of factors to extract. Therefore an cut off value is fixed with an value equal to 1 which can efficiently identify the factors based on eigen values. The following Table 3 signifies the KMO and Bartlett's values as an outcome of factor analysis.

From the tabulated results it has been found that significance level using Bartlett's test is found to be <0.0001 respectively. In Table 4 the variance level has been estimated accordingly with the cumulative value for the 10 components (features) and its varying levels. The percentage of variance is found to be valid with all the 10 components that has been considered for evaluation and its state level estimation strategy.

Table 3. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.863
Bartlett's Test of Sphericity	Approx. Chi-Square	6469.587
	df	465
	Sig.	<.0001

Table 4.Computed Variance and cumulative value

Component	Extraction Sums of Squared Loadings	Rotation Sums of Squared Loadings		
	Cumulative	Total	% of	Cumulative
	%		Variance	%
1	22.553	5.518	17.800	17.800
2	28.756	2.005	6.468	24.268
3	34.155	1.715	5.533	29.801
4	38.756	1.688	5.444	35.245
5	43.086	1.683	5.429	40.674
6	47.300	1.492	4.814	45.488
7	51.319	1.381	4.454	49.942
8	55.029	1.278	4.124	54.066
9	58.586	1.259	4.062	58.128
10	62.011	1.204	3.883	62.011

The following are the attributes (features) that has been selected accordingly with the evaluation through the hybrid approach.

1. Cumulative average GPA
2. Time Spend Learning_AC
3. Distraction_AC
4. FixedBedTime_AC
5. AffectSleepingHabits_BC
6. U_StrengthenSocialPersonality
7. StayingHome_Laziness
8. ELearning_unhealthy
9. TestinHome_notComf
10. Lockdown_Stress

With the above set of attributes and its splitting values we have identified the section-wise attributes as in Table 5 that significantly corresponds to the mental health interventions.

Table 5.Section-wise attributes and its level

S.No	Section	Attributes
1.	Use of digital tools	Cumulative average GPA Time Spend learning_AC Distraction_AC
2.	Sleeping habits	FixedBedTime_AC AffectSleepingHabits_BC
3.	Social interaction	U_StrengthenSocialPersonality StayingHome_Laziness
4.	Psychological state	ELearning_unhealthy Lockdown_Stress
5.	Academic performance	TestinHome_notComf

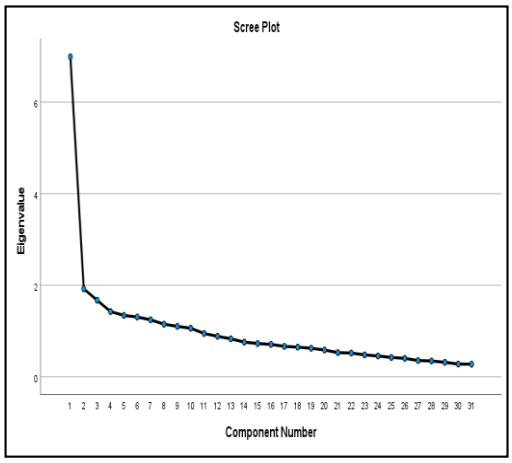

Fig .5 Scree Plot Analyses

From the selected set of features an plot analysis is made to determine the correlative nature among them. The axes have been labelled among the attributes component and the eigen value of each of the component features. From the curve it has been observed that when the component number increases then the plot significantly decreases with regard to the eigen value of each of the component.

6. Conclusion

This research analyzed the e-mental health interventions in virtual learning using hybridization of genetic algorithm with an artificial neural network for the data analyzed from Jordanian University. The experiments evaluated that behaviourism is found to be a key factor which made a significant variation in academic performance using ICT in engineering education and student's learning behavior. The e-learning health interventions corresponding to time spend in learning, distraction, laziness, conditions of e-learning making unhealthy, and the stress in lock down situation made the students to have a diversion from learning and understanding. With the proposed model the situations in e-learning can be easily observed and analyzed accordingly. Also, the relevance among the selected attributes. In this research work we discussed about the e-mental health interventions and the relationship that exists among the attributes. The incorporation of factor analysis best describes the variability, among the correlated attributes and helps to determine the joint variations for the unobserved latent level variables. We have formulated a mechanism for the best determination of factors that contribute towards virtual learning environment. This phenomenon is understood in advance, can make the learning environment user friendly and interactive among the student's community.

References

1. Jae-Eun Russell, Anna Smith and Russell Larsen (2020) Elements of Success:Supporting at-risk student resilience through learning analytics, Computer & Education, 152, 1-14.
2. Vidyut Rajhans, Usman Memon, Vidula Patil and Aditya Goyal (2020) Impact of COVID-19 on academic activities and way forward in Indian Optometry, Journal of Optometry, 13, 216-226.
3. Akhila Joshi, Padmashree Desaiand and Prakash Tewari (2020) Learning Analytics framework for measuring students' performance and teachers' involvement through problem based learning in engineering education, Procedia Computer Science, 172, 954-959.
4. Juan L.Rastrollo-Guerrero, Juan A. Gomez - Pulido and Artuo Duran-Dominguez (2019) Analyzing and Predicting students' performance by means of Machine Learning:A review, applied sciences, 10, 1-16.
5. Bolinskia .F, Boumparisa.N, Kleiboer.A, Cuijpers.P , Eberta.D.D and Ripera.H (2020) The effect of e-mental health interventions on academic performance in university and college students: A meta-analysis of randomized controlled trials, Internet Interventions, 20, 1-10.
6. Lokanath Mishra, TusharGupta and AbhaShree (2020) Online Teaching-Learning in Higher Education during Lockdown Period of COVID-19 Pandemic ,International Journal of Educational Research Open,1-25.
7. FanYang and Frederick W.B.Li (2018) Study on student performance estimation, student progress analysis, and student potential prediction based on data mining,.Computers & Education, 123, 97-108.
8. Alberto Rivas, Alfonso Gonzalez-Briones, Guillermo Hernandez, Javier Prieto and Pablo Chamoso (2020) Artificial neural network analysis of the academic performance of students in virtual learning environments, Neurocomputing, 1-8.
9. Nikkie Zanevsky, (2016) Four critical questions to help you set your student learning objectives, https://www.nwea.org/blog/2015/four-critical-questions-to-help-you-set-your-student-learning-objectives/.
10. Ouafae, E, & Yasser, EI, & Lahcen, O, & Ahmed, D. (2019) A Multiple Linear Regression-Based Approach to Predict Student Performance, Advanced Intelligent Systems for Sustainable Development (AI2SD'2019), pp.9-23.
11. Sheik Abdullah, A, Gayathri, N, Abirami, AM Parkavi, R, (2018) A Statistical Approach for setting SLO Targets over Outcome Based Education-A Case Study, Journal of Engineering

Education Transformations, DOI: 10.16920/jeet/2018/v0i0/120898

12. Tai, Y., & Ting, Y.-L. (2019). English -learning mobile app designing for engineering students' cross-disciplinary learning and collaboration. Australasian Journal of Educational Technology, 120–136. doi:10.14742/ajet.4999.
13. Parkavi, R, & Karthikeyan, P, & Sheik Abdullah, A. (2018) Impact of Massive Open Online Courses and Best Practices: A Case Study on Social Network Analysis Course, Journal of Engineering Education Transformations, Vol.31,No.3, DOI: 10.16920/jeet/2018/v31i3/120782.
14. Parkavi, R., Sheik Abdullah, A., Sujitha, S., Karthikeyan, P, (2017), 'Mobile Devices in Class Room', Handbook of Research on Mobile Devices and Smart Gadgets in K-12 Education, IGI Global Book Publishers.
15. Larson J., Jordan S. S., Lande M. and Weiner S., (2020) "Supporting Self-Directed Learning in a Project-Based Embedded Systems Design Course," in IEEE Transactions on Education, vol. 63, no. 2, pp. 88-97, doi: 10.1109/TE.2020.2975358.
16. Lekha Das , & Megha Sharma (2020) Case Study: Use of MOODLE to Improve Teaching Learning Process in Control System Engineering Vol. 33, No. 3, DOI: 10.16920/jeet/2020/v33i3/144640.
17. Levine Laufgraben, (2016) Using assessment results: setting benchmarks, Faculty Senate Assessment Committee, MSCHE Workshop.
18. Baker, G.R., Jankowski, N., Provesiz, S. and Kinzie, J. (2012). Using Assessment Results: Promising Practices of Institutions That Do It Well. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).
19. Duarte A. J. et al., (2020) "Engineering Education for Sustainable Development: The European Project Semester Approach," in IEEE Transactions on Education, vol. 63, no. 2, pp. 108-117, doi: 10.1109/TE.2019.2926944.