ISSN: 0974-5823 Vol. 7 No. 1 January, 2022

International Journal of Mechanical Engineering

MACHINE LEARNING APPROACH TO AUGMENT PERFORMANCE OF ISED LEVEL-1 STUDENTS THROUGH THEIR ONLINE LEARNING BEHAVIOUR

Somdeep Das[1], Dr. Pinaki Pratim Acharjya[2], Dr. Hiranmoy Mondal[3], Dr. Mauparna Nandan[4]
1Department of Computational Science, Brainware University, India
2 Department of Computer Science and Engineering, Haldia Institute of Technology, Haldia, India
3 Department of Mathematics, MAKAUT, India

4 Department of Computational Science, Brainware University India

Abstract

Online learning has risen exponentially in higher education over the recent times, but it has been confined to a small group of students. However, the Covid 19 pandemic, on the other hand, has culminated in an extraordinary expansion in the online learning environment due to the closure of schools, universities, and institutions. This has undoubtedly been one method of mitigating the difficulty throughout the last year of present worldwide calamity. The advent of internet and communication technologies has undoubtedly been the foundation of these Online Digital Learning Systems (ODLS), where students and teachers may connect to a single platform for purposeful learning. This paper presents the study and analysis of Machine Learning strategy to improve the performance of ISED Level 1 students (6-12 years) through their online learning behaviour. In this methodology, various Machine learning algorithms are used for the analysis of student's online learning behavior in ODLS and help teachers to adopt appropriate approaches towards each student with the predicted results. Firstly, the K-means algorithm is used for clustering the dataset and subsequently, machine learning classifier algorithms are used to construct the predictive model.The accuracy and F1 score of KNN algorithm was observed as the best among the other algorithms with F1 =0.99 and accuracy of 99%. This study also put forth the significance of designing a holistic learning environment for future, specifically the cognitive and emotional engagement of students of ISED Level-1 students, whose behavioral aspects varies with their varied levels of education.

Keywords: Machine learning, Online Learning, Supervised ML Algorithms, Clustering, Classification.

1. Introduction

In the recent times technology has transformed the education system, the way students learn, and the teachers teach. The online digital learning system (ODLS) is the trend which is evolving rapidly over the traditional learning systems. The success of digital learning system, which is enhanced and powered by technology relies on the digital interfaces/platforms, the availability of internet connectivity, internet-enabled devices, and the readiness of learners and teachers.

Online teaching and online delivery of courses have become the main advantages of ODLS as it is leading to a revolution by exposing the world-class education system to the students. Unlike traditional systems, instead of physical classrooms, virtual classrooms make the platform for teaching and learning environment. Virtual classrooms provide the flexibility which is beyond time and space.As a result, majority of students can register themselves inthese online curriculums [4]. Through the integrated software portals, students can track their performance and interact with their teachers. In the similar way the parents can also track the performance, behavior, and attendance of their wards and interact with their teachers. The online pedagogical system provides a significant paradigm shift in the education system and has brought student, parent and teacher together at one platform.

Key Users of Digital Platform: Teachers from Kindergarten to ISCED level-3, College Professors and Students, and Tutors.

1.1 The digital interface/ platform: Digital learning platforms provides with an interactive interface for students to explore lesson content in virtual environment facilitated by digital equipment. Here, digital equipment can refer to software applications and devices both. These platforms essentially leverage lessons with significantly amalgamate with multimedia contents, videos, animations, photos, and audio recordings. These platforms provideself-evaluation basis learning which further enables students with an opportunity to do and redo their activity to achieve excellence. Digital learning platforms have the potential to be an effective learning tool for students at many levels of education.Hence the key benefits of these interfaces can be summarized as:

- Multimedia learning experiences for students
- Differentiate instruction automatically based on student ability
- Self-evaluation of students' progress
- Virtual Exercises
- No barrier of Physical Classroom
- Student-Teacher-Parent involvement
- Cheaper

Digital Learning Platforms integrating with the Software and Services like various management and information systems along with document creation software builds a robust online environment for learning. LMSs (Learning Management Systems)work with other educational platforms, like SISs (Student Information Systems) and other document creation software which further provide the virtual learning environment for students and teachers. While SISs digitally maintain the documentation part of the student information, LMS performs the dissemination of the courses through the online platform [2, 3]. Whereas Document creation software enable instructors to create their own lessons often allow the user to upload files from document creation software.

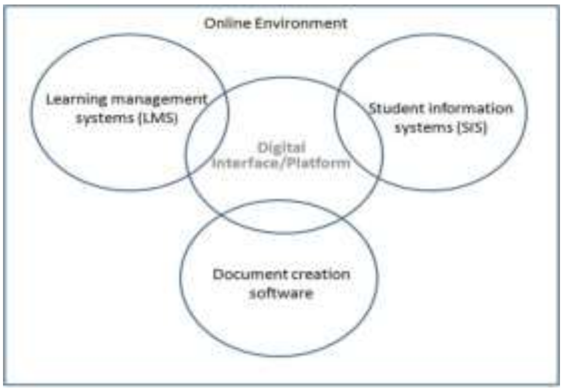


Fig.1 Shows Digital Interface integrated with software and services in an Online Environment

The education progression has been immensely transformed by the advent of innovative digital tools, applications, and devices. In the learner's perspective the evolution of ed-tech sector has made the delivery process to be accessed from everyplace and anytime. These ed-tech sectorsare more engaging, entertaining and exploring. With the arrival of 4G internet at affordable rates, and affordable Smartphone access to technology to retrieve course content has profoundly impacted the online digital courses. The leading population of new learners are accessing internet from their Smartphone which is a perfect, customized and business empowered stage for online training.Further, various online Education Technology applications (Edtech), Massive Open Online Courses (MOOCs)[3,5], Web based education platforms, and cloud-based communication platforms can be easily accessible on Smart phones besides tablets and laptops. The interpretation of statistic [33] shows that China has more than 850 million users of smartphones as represented in (Fig.2) and which is more than any nation in the world. In terms of more users of smartphones, India stands second in the ranking, but not exactly half the same number of as China. The figure (Fig.2) shows the year wise growth of smartphone usersby country.

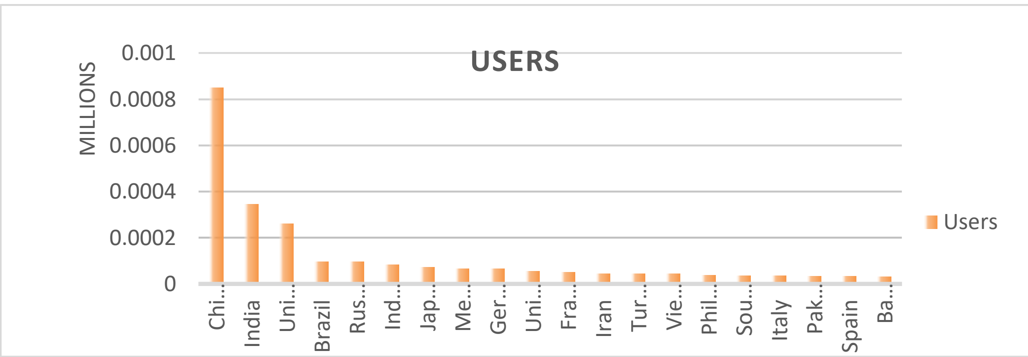


Fig.2 Shows high and increasing Smartphone penetration

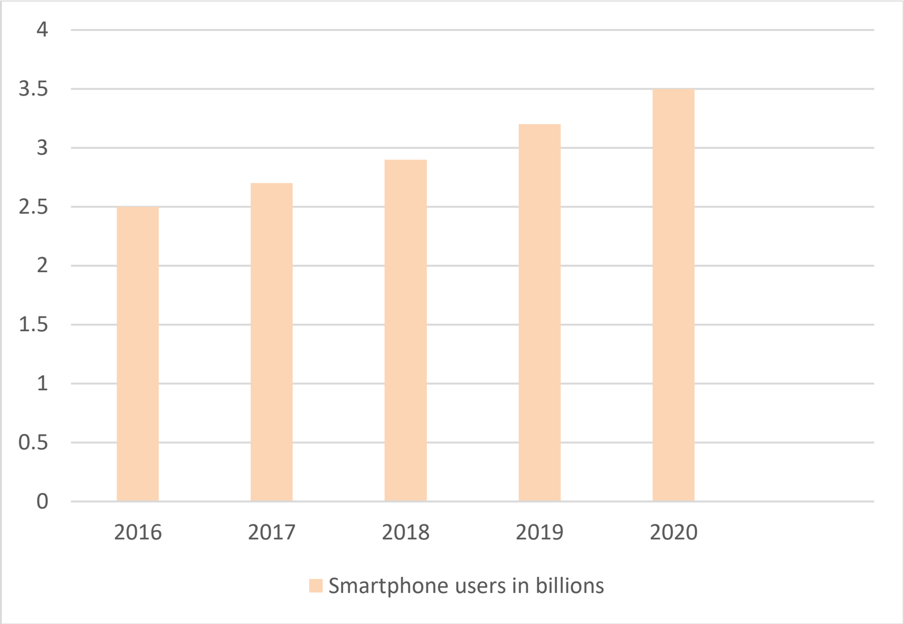


Fig.3 Shows year wise increase in Smartphone users

The significance and popularity of ODLS can be visualized in two phases, first pre-Covid19 and post Covid19 outbreak. Before the pandemic hit the global scenario resulting in shutdown of schools, colleges and Universities to contain the outbreak of Covid19, ODLS was gaining pace and popularity for self-motivate learners, executive learners and tertiary education level students for honing skills in specific courses. Whereas Digital Learning Systems like Smart Classroom, Flipped Classroom [16], Use of Video etc. were moderately popular in primary and secondary level schools.

Post Covod19, the scenario has changed. ODLS has become a necessity for students from Primary level education to professional level learning. The traditional physical classroom curriculums are converted into online courses, and virtual classroom solutions to allow millions of students to continue learning from their homes. The online education is transforming the domain of education, connecting students to global learning platforms, and making learning more dynamic. Apart from this, few decisive factors affecting the students were also observed in online learning like motivational factors to learn, factors related to the learning platform/ interface and the factors related to the new teaching and learning procedures [6]. Different cloud based online platform like Zoom, Webex and so forth integrated with LMS like open-source learning management frameworks like Moodle are developing as virtual learning foundation of ODLS, though Google classroom [34] itself felicitates the services of online learning management system.

Increasing internet penetration, time constraints faced by students, geographical challenges in attending physical classes, and the low expense involved in online training are the primary drivers of the digital learning sector. Going forward, there would be a convergence of digital and physical worlds, as internet becomes more flexible and accessible to everyone.

In the backdrop of Covid-19, while these ODLS are gaining popularity, certain challenges are also associated with online learning systems [3, 4] as there is no direct communication between student and teacher. Dearth of student motivation and related concern in various course activities and course information makes the most critical challenges of this learning environment. With the emergent of internet-based learning platforms it has become easier to monitor student performance by the teachers. The log data can be used to study and analyze the performances of students. Learning Management Systems (LMSs), Massive Open Online Courses (MOOC), and Digital Electronics Education and Design Suite (DEEDS) [4]are some of the most looked afterweb-based learning platforms. Albeit various models were verified based on MOOC, most of them covered only the behavioral aspect of learning engagement, as opposed to the multidimensional parts of association in learning [11]. However, research has shown that there is an interrelation amongeducational data mining, psychology, and education. Ithas established that students' active involvement and approach considerably impact learning [19].

Many studies are carried out to predict the students’ performance on their conduct in the classroom. Unlike [6], this states that strength of relationship between student engagement and their respective performance in online course is positive but weakly related. Predictive analytics, helps instructors to monitor students’behavior in terms of utilization level of course contents and student’s evaluation scores as they aredirectly associated with student’s involvement level [3]. On the other hand,analytics facilities, the students by providing them suitable suggestions and solutions to enhance the learning performance by collecting data and further analyzing the collected data[3, 9].

1.2 In this research a variety of machine learning algorithms are used for prognostic analysis of the student's beahviour under certain conditions and building the model. The presented study focuses in building a machine learning model to comprehend the student's behavior and classification under ODLS.

The works related to the use of Machine Learning is to help the teacher to generate better models and learning methodologies applied in the online digital learning environments. The analysis result allows the teachers to avail preemptive solutions for their teaching approaches towards the students. Various LMSs are facilitating in capturing this data for the teachers.

The principal goal of the ML approach is to create a model that may be used to carry out classification, prediction, estimation, or similar comparable tasks. The purpose of prediction is to evaluate the unknown value of a variable based on relevant historical data.

In the present work, this variable is related to the online learning behaviour of ISED Lavel-1 students and their respective performance in the class. In other words, the proposed assessment or prediction of student performance is based on a number of historical academic characteristics that describe student behavior.

2. Background Study and Mapping of Features

To understand the impact of ODLS on the behavior of school level students we need to first classify the various age group of students ranging from 6-18 years old in an Internationally recognized standard and then we need to understand the psychological aspects of those classified age groups with respect to their memory processes, cognitive development and**,** maturity levels. After classifying the different age group levels, prognostic analysis can be done on each of them using Supervised Algorithm to understand the psychological behavior towards ODLS and the of behavioral impact on their performance.

2.1. International classification of education levels:

ISCED is intended to fill in as a system to categorize educational activity as characterized in curriculum and the subsequent capabilities into universally concurred classifications. The fundamental ideas and meanings of ISCED are in this manner expected to be universally substantial and complete of the full scope of instruction frameworks**.**“The International Standard Classification of Education” (ISCED) is a part of the “United Nations International Family of Economic and Social Classifications”and received officially by the General Conference of UNESCO Member States [25, 30]. ISCED is considered to be the reference standard for classification of education curriculums by various education levels and age groups of students.

Level	**School**	**Grades**	**Age range**	**Duration (yrs)**
ISCED Level-0	Early Childhood Development		0-3	
	Pre-Primary Education	KG	3-6	
ISCED Level-1	Primary education	Grade-1	6-7	**6**
		Grade-2	7-8	
		Grade-3	8-9	
		Grade-4	9-10	
		Grade-5	10-11	
		Grade-6	11-12	
ISCED Level-2	Lower Secondary	Grade-7	12-13	**2-5**
		Grade-8	13-14	
		Grade-9	14-15	
		Grade-10	15-16	
ISCED Level-3	Upper Secondary	Grade-11	16-17	**3**
		Grade-12	17-18	
ISCED Level-4	Post-Secondary	**-**	**18-21**	**2-3**
ISCED Level-5	Short cycle Tertiary education	**-**	**18-20**	**2**
ISCED Level-6	Bachelor’s or equivalent	**-**	**18-21**	**3-4**
ISCED Level-7	Master’s or equivalent	**-**	**21-23**	**2**

Table 1: Classification of various ISCED Levels

Table.1, summarizes various levels of education system spanning from ISCED Level-0 to ISCED Level-7 with respective age range, grades and duration taken to complete the level.

We consider the ISCED Level (1-3) in our study, since this belongs to the pure School level classification. With this we can classify the basic education intotwelve years of duration which includes aPrimary Education of six years duration, Lower Secondary Education of four years of duration, and Upper Secondary Education of two years duration.

The former ISCED level (0) has learning pedagogy planned for younger children ranging in the age group of 0 to 6 years, whileanother is planned for ISCED levels (5-8).

2.2. Behavioral phenomena:

To further justify the psychological behavior of these age groups as mentioned in the “Primary”, “Lower”“Secondary”, and “Upper Secondary” level education system, we need to understand the various stages involved in the cognitive development and learning with respect to cognitive and educational psychology.
The cognitive stage theorystates the gradual phase by phase process of the ability to think rationally and methodically [35]. Cognitive development refers to the transformation in the progression whereas cognition alludes to perception and processing of memory [27].

Stage No	**Stage**	**Ability**	**Age group (in years)**
1	Sensorimotor	“Ability to experience the world through movement and senses.”[35]	Birth to age 2
2	Preoperational	“Ability to symbolize activities, but the activities are considered logically inadequate” [35]	2 to 7
3	Concrete Operational	“Ability to think and understand more flexibly and logically but not up to the standard of an adult” [35]	7 to 11
4	Formal Operational	“Ability to think rationally and hypothetically and develop problem solving skills” [35]	11 and beyond

Table 2: Summary of Jean Piaget Theory

A theory developed by Jean Piaget on cognitive development illustrates the transformation in logical thinking which occurs during the growth ages of a child. As presented in Table.2, the theory, further explains four phases of cognitive development [27]and summaries the thinking, learning, and reasoning ability of a child with respect to their age groups. It also shows the how the ability of child vary from one phase to another in terms of cognitive development.

Psychologically,behaviour of student is one of the root drivers for their performance at a school.All psychological behavioural issues like self-motivation and self-driven attitude add to learning and it is not merely intelligence that influences learning outcomes [22]. Also, the learning capability of a child variesfrom one child to another. This ambiguity also implies on the development factor of a child. Apart from this, the cognitive development of a child is also highly influenced by the social aspect of their ecosystem. Child’s interactions with their parents, teachers and peers greatly impact their intensity of thoughts and understanding [26]. In psychological science, well-defined and established standards are used for appraisal processes to measure the students’ performances [36]. A learning appraisal is often the initial phase in this progression.

As stated in [29], learning appraisal reveals the following outcomes:

- Comparison of student’s age and education classification level to achieve academic success
- Strengths and weaknesses in hidden learning skills
- Positive and negative motivation
- Environmental concerns

In the online platform the behavioural aspect can be understood by the level of participation and involvement in the class and respective learning outcomes.

3. Related Work

Observation of researchers depicts that the correlation between behavioral patterns and academic performance are very strong [17]. In pursuance to analyzing the psychological behavior of students included in the ODLS environment, it was discovered that considerable research on predicting various behavioral patterns of students has been conducted.

C et al. states that use of both behavioural and learning data offers better results used by the Support Vector Machine prediction algorithm whereas diminishing the number of features does not optimize the outcome of prediction. The results showed that, in a learning environment, behavioural and learning data provides better and more stable predictions about the results of the students. It concluded that the performance of the predicting algorithm is augmented by the use of assorted significant data. [2]

The study made by Liu and d'Aquin presents the relationship between student characteristics and their learning achievement. In this approach they used supervised learning algorithm to predict the learning achievement of student included in the virtual learning environment. Moreover, they used K-Prototypes algorithm with different K values on different set of experiments to find the cluster of successful and unsuccessful students. They concluded that successful students are more mature and active, their upbringing is in privileged areas, and have completed their higher education as compared to the weaker counterparts [21].

F et al. in their study suggested course précised predicting models to classify weak students so as to help them get better in their performance. "Support Vector Machine", "K-Nearest Neighbors", and "Naive Bayes Classifier" are the three-machine learning supervised algorithms, which were used in this study to identify the best algorithm. Among the various methods used for testing, the Naive Bayes Classifier model's performance was identified as the best. It concluded that by providing specific guidelines on creating accurate prediction model to identify weak students, the instructors and the students both can be informedin advance to improve the performance of students. However, it was found in the study that pedagogical decisions made by the teachers such aspedagogical design, assessments process, and the grading system instructions tend to limit this model on these circumstantial factors [23].

M et al. explored the possibility of incorporating various machine learning algorithms such as an "Artificial Neural Network" (ANN), "Support Vector Machine" (SVM), "Logistic Regression" (LR), "Naïve Bayes Classifier"(NBC), and "Decision Tree"(DT). The result showed that among the above algorithms, ANN and SVM models achieved the precision of 75% in forecasting the course related complexity which a student will face in the next session. Subsequently all the models in the study were trained and found that an SVM model achieved an accuracy of up to 80%. The researchers observed that using an SVM or ANN is suitable for forecasting the performance of an individual student and can improve the performance of teaching, learning and student involvement. They observed that studied model can identify weak students in advance and teachers can revisit designing of those digital courses for them. Such augmentation can guide teachers to identify the session complexity in advance and hence anticipate the learning behaviors of student during various coursework and exercises [4].

Tomas and Jayagopianalyzed the attentiveness of the students by capturing data points of various facial gestures of students. Computer vision techniques have been used for capturing the data points. Further, those data points were fed into various supervised machine learning algorithms for classification and the results were tabulated. The hyper-parameter of the final model was cross validated to address the common issue of over fitting in supervised learning algorithms. "Support Vector Machine" and "Logistic Regression" algorithms were used to generate the final models. They proposed a predictive model of student's engagement in a classroom. Applying the model on the video data recorded from the classroom illustrates the state of students in terms of their being focused or unfocused and subsequently the model makes decision. They also compared machine learning algorithms and baseline evaluator to conclude that the performance of machine learning algorithms is better as compared to the later [8].

M et al. qualified several Machine Learning algorithms to develop a model to test the result for predictingspecific behavior of students on low-engagement. The results demonstrated that in terms of accuracy the "J48", "Decision Tree" (DT), "JRIP", and gradient-boosted classifiers displayed better performance as compared to other tested models. On the basis of those obtained values, they conducted two experiments, one of which illustrates that DT, J48, JRIP, and GBT are the largelysuitable algorithms for predicting low engagement students. Another experiment for predicting low-engagement showed that the most important input variables of the dataset are clicks on specific tabs of website. The researchers concluded that proposed predictive model allowteachers to engage and ensue students in different activities [3].

H et al. identified three characteristics of student viz. diligence, orderliness and sleep patterns, which have significant correlations with academic performance. The researchers calculated the result on prediction by grouping various features which are related to academic performance. They designed and proposed a "Multi-Task Learning-to-Rank Academic Performance Prediction framework" (MTLTR-APP). Various algorithms like Ridge Regression, Decision Tree, Random Forest, RankSVM, MLP, and XGBwere used and found that MTLTR-APP algorithm can perform better in predicting academic performance [17].

K et al. proposed the use of probabilistic graphical models, which significantly improve prediction accuracy. The researchers also presented student knowledge and examination strategies jointly in one model. Furthermore, they concluded that as opposed to pure performance models, enhanced models perform better results for learning [19]. Ding et al. [7] explored the intensity of student involvementin online forums and the impact of 'gamification' on them.

K et al. examined the qualities of students that impact the learning conduct and for cleaning the original dataset used the attribute reduction method. They ascertained the semblance of understudies' conduct and used the Jaccard coefficient calculation to characterize the understudies. Furthermore, they utilized three classifier models as Linear Discriminant Analysis,"Logistic Regression, and Linear Support Vector Machine[28].

Although most of the literatures were found to be focused on higher education, the respective predictive model obtained from the behavioral phenomena observed in colleges, universities, and tertiary education can possibly show the way forward for the study and analysis on student behavior in ISCED Level-1 (i.e. Primary Education). Various behavioral patterns impacting academic performance has been identified [22] such as student's active participation in the class, raising hands for queries or answers, involvement in discussion, regularly visiting the posts, and presence in class. These patterns can be classified as the key attributes of the study.

4. Methodology

Our aim is to analyze behavioural patterns of ISCED level-1 students in an online digital learning environment. Public domain dataset [31] has been taken. Data preprocessing and analyses is performed using Python 3.7.6.

4.1 Data preprocessing

Pre-processing plays an important role in Machine learning. The main purpose of preprocessing is to convert the raw dataset into suitable form so that we can find the most suitable input variables which aremostly connected to the students' behavioural aspects. As suggested in [22], we can prepare our dataset in the category of age group of 6-12. Also, from [27], we can classify this age group in ISCED level-1. The dataset does not consist of any attribute mentioning the age group. Hence, the dataset is prepared by grouping the Grade attribute from Grade-1 to Grade-6 (age group of 1-12). First, this prepares the dataset into our desired ISCED level-1 age group of students, which also defines the Primary Education students' age group. Student performance is another feature which is understood by the Grade classification attribute of the dataset. Different Grade classification levels and their respective values are Low-Grade (0 to 69), Middle-Grade (70 to 89), High-Grade (90 to 100). Further, these levels are represented as L, M, and H respectively. These preprocessing does not imply any logical change in the dataset.

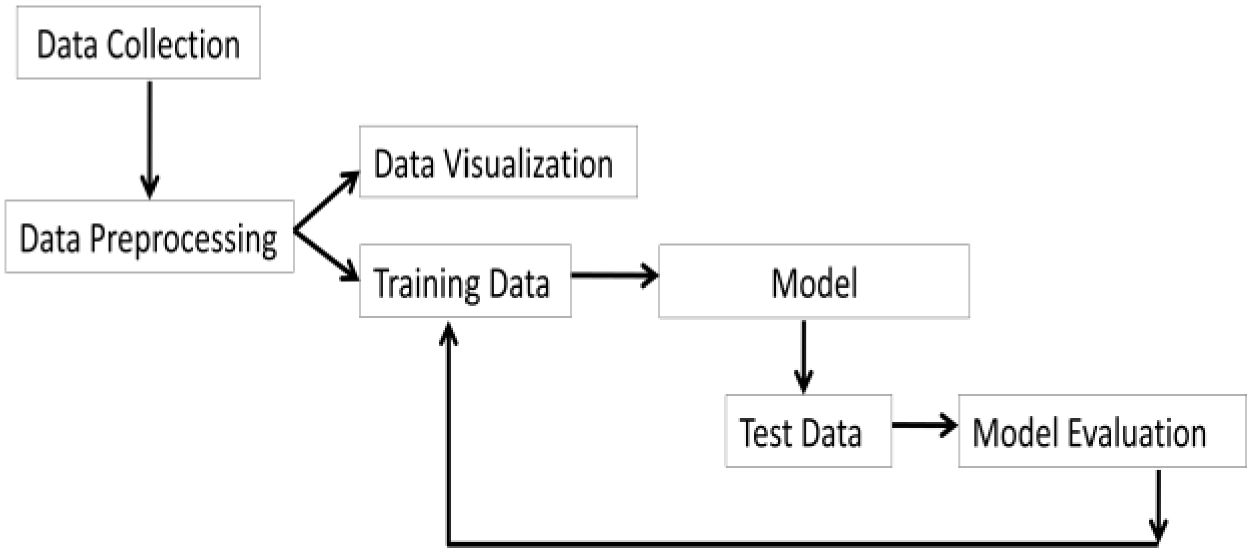


Fig. 4 Schematic diagram of the prognostic model

4.2 Feature Selection

Feature selection is a method whichselects an appropriate subset of attributes from the original attribute which can proficientlydepict the input variables [4]. The performance of the learning model depends on the input variable i.e. the feature selection. In our dataset there are many attributes which may beirrelevant for classification purposes and may impact the model negatively. This issue can be resolved by selecting input variable from the dataset which may describe the student's characteristics appropriately and further can be utilized to predict the performance of student.

In our study we extracted few those features which may impact in predicting students' performance. We extracted features (Raised hand, Visited resources, Viewing announcements, and Discussion) from the dataset, which may be termed as the behavioural features. This seems relevant for understanding the students' engagement during the class. To do this feature selection we used Exploratory Data Analysis (EDA). Other featuresmay also be separately used to understand the impact on students' performance.

4.3 Visualization

Data visualization primarily gives an idea of the data and shows different parameters and provide qualitative understanding of the dataset and can help in identifying various patterns and relationship against each other. Here, in Fig. 5 we have used the histogram to show each of the identified parameters from the preprocessed dataset to indicate the behavioural features of students. The scattered matrix presented in Fig.6 shows statistical overview of differentbehavioural input parameters against each other.

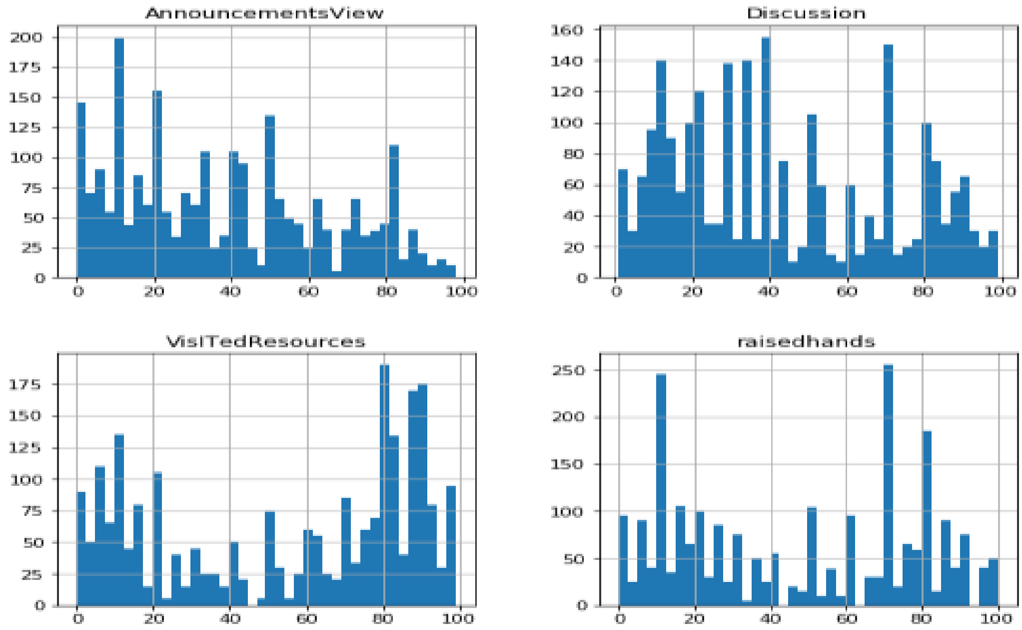


Fig. 5 Histogram showing each identified beahaviour values

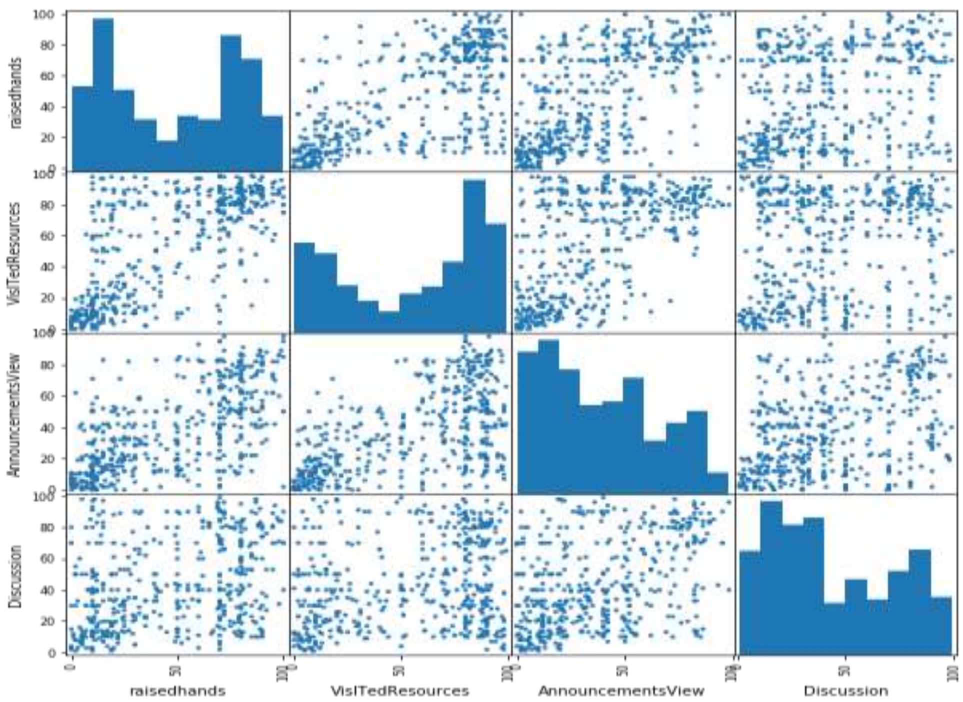


Fig.6 Scattered Plot showing different behavioural inputs

We can make predictions about one variable from another if there is a relationship between two variables. Further toidentify the strength of association between two features of behaviour, Pearson *r* correlation is applied between the studentsbehaviour.It shows how the features are correlated to oneanother. Pearson *r* correlation is the most popular types correlation, used to compute the degree of the relationship between linearly associated variables [13]. The following is the formula to calculate the Pearson *r* correlation:

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

r_{xy} = Pearson r correlation coefficient between x and y
n = number of observations
x_i = value of x (forith observation)
y_i = value of y forithobservation)

In terms of the strength of relationship, a correlation of –1 indicates negative correlation. A correlation of +1 indicates positive correlation. A weaker relationship between two variables is depicted by the correlation coefficient value tending to 0.

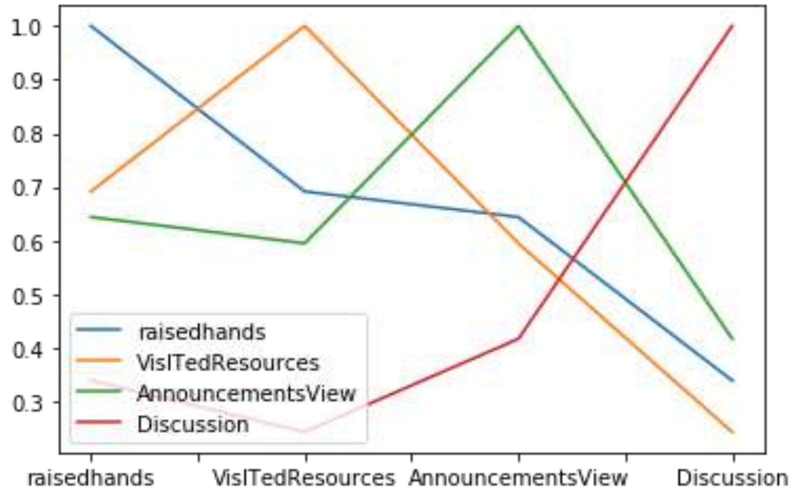


Fig.7 Correlation graph of features

	raisedhands	VisITedResources	AnnouncementsView	Discussion
raisedhands	1.000000	0.639584	0.587896	0.234832
VisITedResources	0.639584	1.000000	0.518894	0.155941
AnnouncementsView	0.587896	0.518894	1.000000	0.426953
Discussion	0.234832	0.155941	0.426953	1.000000

Table 3 Correlation of features

Applying Pearson r correlation, it is clear from fig. 7 and Table. 3, the best correlated features are raisedHands and visitedResources.
Furthermore, a K-means algorithm is applied on a new dataset extracted with the behavioural features from the original dataset to create the cluster of student. An elbow method is used to find the best k, as k=7 for K-means algorithm. The student clustering is done in three groups Low, Mid and High. Figure 8 shows the scatter plot of students clustering. From this visualization, it can be observed that there are three clusters represented by 3 colours. Each cluster has its own centroid.

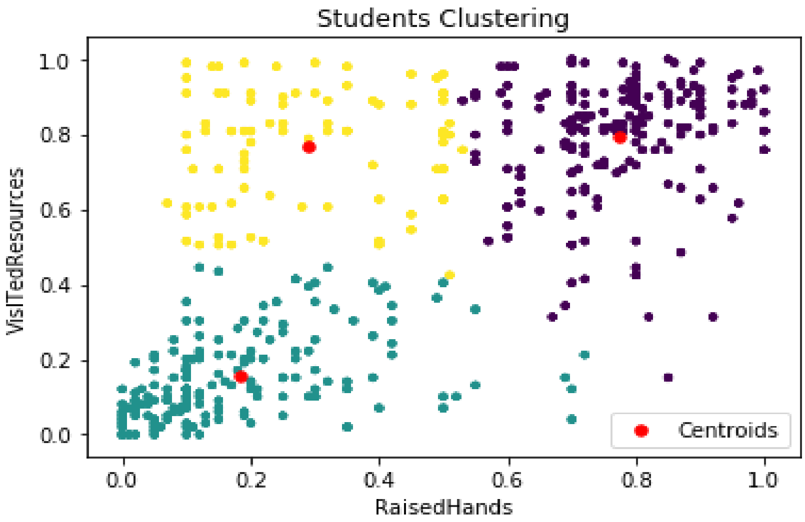


Fig.8 Clustering of students with behaviour

4.4. Model Construction

To predict each student's behavior and respective performance in an ODLS, we use several widely used machine learning algorithms, viz. K-Nearest Neighbours, Logistic Regression, Support Vector Machine and Decision Tree.Based on the performance of the classifier algorithms we can suggest the best model for this study.SVM is a supervised learning method which can perform both linear classification as well as non-linear classification unlike LR which is only used for binary classification [8].DT is a popular algorithm for classification as well as regression whereas KNN is a classification algorithm generally used to predict categorical value.

Now we first split our preprocessed data in the training dataset and testing dataset. Then input the training dataset to each of the algorithms to train the model and hence constructing the model.Thereafter, the defined model is tested with the test dataset for attaining the utmost accuracy. The predictive model is then evaluated with evaluation matrixes for maximum accuracy and will be deplyod in the proposed system.

4.5. Model Evaluation

To quantify the accuracy of the algorithm we use confusion matrix. The confusion matrix is used to illustrate the performance of each classifier algorithm applied on a group of test-data for which true values are known. By performing this we put the model into its performance evaluation state. Confusion matrix provides the quantitative information with four values, TP (True positive), FP (False positive), FN (False Negative), and TN (True Negative). All the number of values (TP+TN+FP+FN) in the matrix represents the number of data in the test data. Fig. 9 shows the formation of a confusion matrix. These four parameters of performance evaluation of a model i.e.accuracy, precision, recall, and F1 score are used to evaluate the effectiveness of a model.

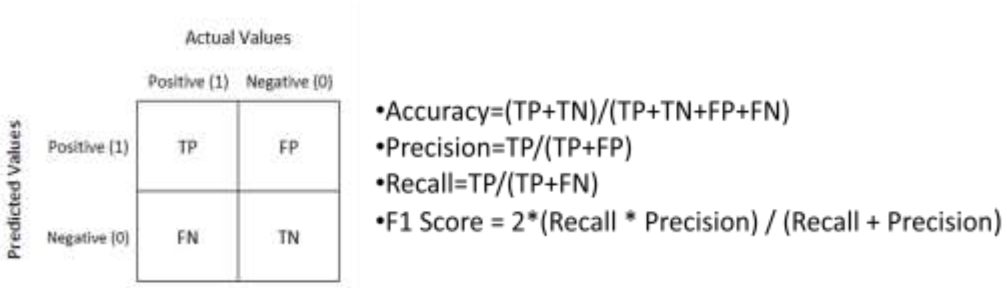
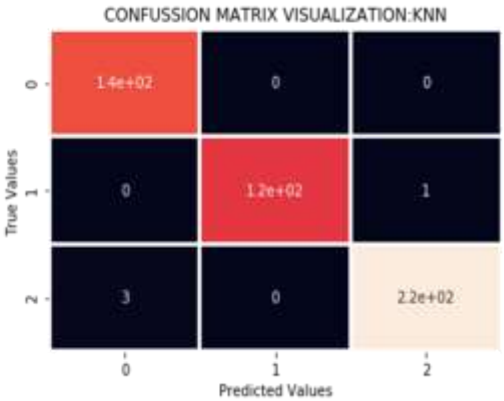


Fig.9 A simple confusion matrix construct

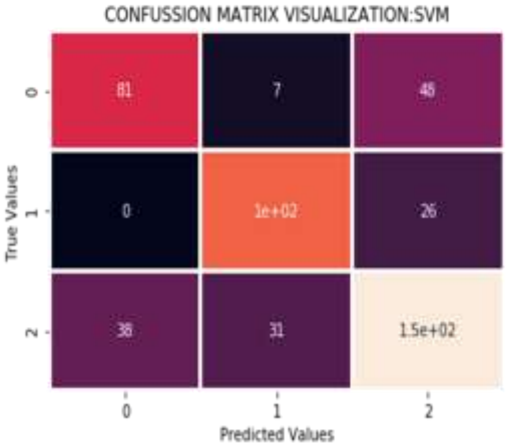
To evaluate the performance of our models we find the confusion matrix for each of the tested models built with KNN, SVM, LR, and DT. Fig. 10, 11, 12 and 13 shows the confusion matrixes of each of the performing algorithms. Also it shows the respective values of performance evaluation parameters accuracy, precision, recall, and F1 score.



KNN:

	precision	recall	f1-score	support
H	0.98	1.00	0.99	136
L	0.87	0.99	1.00	126
M	0.44	0.99	0.99	218
accuracy			0.99	480
macro avg	0.99	0.99	0.99	480
weighted avg	0.99	0.99	0.99	480

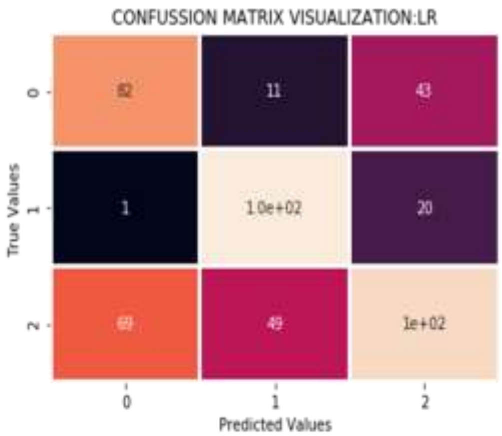
Fig. 10 Confusion Matrix of KNN Model and Performance Chart



SVM:

	precision	recall	f1-score	support
H	0.68	0.60	0.64	136
L	0.87	0.87	0.76	126
M	0.67	0.68	0.68	15
accuracy			0.69	480
macro avg	0.69	0.69	0.69	480
weighted avg	0.69	0.69	0.69	480

Fig 11. Confusion Matrix of SVM Model and Performance Chart



LR:

	precision	recall	f1-score	support
H	0.54	0.60	0.57	136
L	0.74	0.93	0.82	126
M	0.50	0.27	0.35	218
accuracy			0.60	480
macro avg	0.60	0.63	0.61	480
weighted avg	0.60	0.60	0.59	480

Fig.12 Confusion Matrix of LR Model and Performance Chart

Fig 13. Confusion Matrix of with DT Model and Performance Chart

Sl No	Algorithms	Accuracy	F1 Score
1	K-Nearest Neighbours (KNN)	0.99	0.99
2	Logistic Regression (LR)	0.60	0.59
3	Support Vector Machine (SVM)	0.69	0.69
4	Decision Tree (DT)	0.80	0.80

Table 4. Summry of the Accuracy of tested Models

Hence from the above table 4, it is found that the best model is with the KNN classifier algorithm having accuracy 0.99.

5. Conclusion

The presented research focused on analysis of psychological behavior of ISCED level-1 students who are using the Online Digital Learning System (ODLS). The study highlights the various means of digital learning systems. It focuses on building a machine learning model to understand the student's behavior and impact on their performance under ODLS environment. The model was trained using various machine learning classifier algorithms. The test results in terms of accuracy suggest KNN as the best model amongst others. The result concludes the accuracy of the KNN model as 0.99.

6. Future Scope

Future work may be carried out for the behavioural study and impact of ODLS on students' performance on **ISC**ED level-2 and **ISC**ED level-3 students. Furthermore, model of this research can be used in order to predict the gender wise behavior aspects of students separately for various levels of education.These outcomes can be improved later on when more data are accessible however they are sufficient to distinguish learning issues and propensities. This prognostic model may also be referred for the analysis of student behavior in the post Covid-19 scenario.

References

1. J. Mutahi, A. Kinai, N. Bore, A. Diriye, and K. Weldemariam,"Studying Engagement and Performance with Learning Technology in an African classroom," in Proceedings of SeventhInternational Learning Analytics & Knowledge Conference,pp. 148–152, Vancouver, BC, Canada, March 2017.
2. C. Villagr ́a-Arnedo, F. J. Gallego-Dura`ın, P. Compañ-Rosi-que, F. Llorens-Largo, and R. Molina-Carmona, "Predicting academic performance from behavioural and learning data,"International Journal of Design & Nature and Ecodynamics,vol. 11, no. 3, pp. 239–249, 201
3. M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, and S. Ali, "Using machine learning to predict student difficulties from learning session data,"Artificial Intelligence Review, pp. 1–27,2018
4. M. Hussain, W. Zhu, W. Zhang, and S. M. R. Abidi, "Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores,"Computational Intelligence and Neuroscience, vol. 2018, pp. 1-21, 2018

5. F.C. Bonafini, C. Chae, E. Park, and K. W. Jablokow, “How much does student engagement with videos and forums in a MOOC affect their achievement?,”Online Learning, vol. 21,no. 4, pp. 223–240, 2017
6. B. Bahati, F. Uno, and M. Tedre, “Can student engagement in online courses predict performance on online knowledge surveys?,”International Journal of Learning,Teaching and Educational Research, vol. 16, no. 3, pp. 73–87,2017
7. L. Ding, E. Er, and M. Orey, “An exploratory study of student engagement in gamified online discussions,”Computers& Education, vol. 120, pp. 213–226, 2018.
8. C.Tomas and D. B. Jayagopi, “Predicting student engagement in classrooms using facial behavioral cues,” in Proceedings of 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education, pp. 33–40, Glasgow,UK, November 2017
9. P. Leitner, M. Khalil, and M. Ebner, “Learning analytics in higher education—a literature review,” in Learning Analytics:Fundaments, Applications, and Trends: A View of the CurrentState of the Art to Enhance E-Learning, A. Peña-Ayala, Ed.,pp. 1–23, Springer, Berlin, Germany, 2017
10. C. Manwaring, R. Larsen, C. R. Graham, C. R. Henrie, andL. R. Halverson, “Investigating student engagement in blended learning settings using experience sampling andstructural equation modelling,”Internet and Higher Education, vol. 35, pp. 21–33, 2017
11. Jung and J. Lee, “Learning engagement and persistence in massive open online courses (MOOCS),”Computers& Education, vol. 122, pp. 9–22, 2018
12. Holmes, “Engaging with assessment: increasing student engagement through continuous assessment,”ActiveLearningin Higher Education, vol. 19, no. 1, pp. 23–34, 2018
13. Nettleton D. Selection of Variables and Factor Derivation. In: Commercial Data Mining [Internet]. Elsevier; 2014. p. 79–104. Available from: https://linkinghub.elsevier.com/retrieve/pii/ B978012416 6028000066
14. "A. Tick "Research on the Digital Learning and E-learning Behaviour and Habits of the Early Z Generation," in Proceedings of IEEE 22nd International Conference on Intelligent Engineering Systems, pp. 33-38 , Las Palmas de Gran Canaria, Spain, June 21-23, 2018"
F. Hegyesi, Gy. Kártyás, J. Gáti, “Answers to the 21st century challenges at a university with technical training,” in the Proseedings of IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY), pp. 365-368, September, 2017.
15. A. Sun and X. Chen, “Online Education and Its Effective Practice: A Research Review,” *JITE:Research*, vol. 15, pp. 157–190, 2016, doi: 10.28945/3502.
16. L. Cheng, A. D. Ritzhaupt, and P. Antonenko, "Effects of the flipped classroom instructional strategy on students' learning outcomes: A meta-analysis," Educational Technology Research & Development, vol. 67 pp.793–824, October, 2018.
17. Huaxiu Yao, DefuLian, Yi Cao, Yifan Wu, and Tao Zhou. 2019. Predicting Academic Performance for College Students: A Campus Behavior Perspective. ACM Trans. Intell. Syst. Technol. 1, 1 (March 2019), 22 pages.https://doi.org/0000001.0000001"
18. Scagnoli, N. I., Choo, J. and Tian, J. (2017), ‘Students’ insights on the use of video lectures in online classes’, British Journal of Educational Technology doi:10.1111/bjet.12572
19. Käser T, Hallinen NR, Schwartz DL (2017) Modeling exploration strategies to predict student performance within a learning environment and beyond. In: 17th International conference on learning analytics and knowledge 2017, pp 31–40. https://doi.org/10.1145/3027385.3027422"
20. Thomas K. F. Chiu, Morris Siu-yung Jong, Ida A. C. Mok2, "Does learner expertise matter when designing emotional multimedia for learners of primary school mathematics?," Education Tech Research Dev (2020), 27 April 2020. https://doi.org/10.1007/s11423-020-09775-4
21. Liu S, d’AquinM(2017) Unsupervised learning for understanding student achievement in a distance learning setting. In: IEEE global engineering education conference (EDUCON), pp 25–28. https://doi.org/10. 1109/EDUCON.2017.7943026"
22. Kassarnig V, Mones E, Bjerre-Nielsen A, Sapiezynski P, Dreyer Lassen D, Lehmann S. Academic performance and behavioral patterns. EPJ Data Sci. 2018 Dec;7(1):10.
23. F, Diefes-Dux HA, Madhavan K (2016) Models for early prediction of at-risk students in a course using standards-based grading. ComputEduc 103:1–15. https://doi.org/10.1016/j.compedu. 2016.09.005
24. N Blundell C, Lee K-T, Nykvist S. Digital Learning in Schools: Conceptualizing the Challenges and Influences on Teacher Practice. JITE:Research. 2016; 15:535–60.
25. Schneider SL. The International Standard Classification of Education 2011. In: Elisabeth Birkelund G, editor. Comparative Social Research [Internet]. Emerald Group Publishing Limited; 2013 p. 365–79. Available from: https://www.emerald.com/insight/content/doi/10.1108/S0195-6310(2013)0000030017/full/html
26. C. Herodotou, B. Rienties, A. Boroowa, Z. Zdrahal, and M. Hlosta. A large-scale implementation of predictive learning analytics in higher education: the teachers’ role and perspective, Education Tech Research Dev, vol. 67, pp. 1273–1306, Oct. 2019.
27. Soran University, Babakr ZH, Mohamedamin P, Kakamad K. Piaget’s Cognitive Developmental Theory: Critical Review. eqr [Internet]. 2019 Sep 30;2(3).Available from: https://www.asianinstituteofresearch.org/EQRarchives/Piaget%E2%80%99s-Cognitive-Developmental-Theory%3A-Critical-Review-
28. Liang K, Zhang Y, He Y, Zhou Y, Tan W, Li X. Online Behavior Analysis-Based Student Profile for Intelligent E-Learning. Journal of Electrical and Computer Engineering. 2017;2017:1–7.

29. Livieris IE, Drakopoulou K, Tampakas VT, Mikropoulos TA, Pintelas P. Predicting Secondary School Students' Performance Utilizing a Semi-supervised Learning Approach. Journal of Educational Computing Research. 2019 Apr;57(2):448–70.
29. https://www.melbournechildpsychology.com.au/blog/educational-psychology-how-it-helps-learning, accessed on April, 21st, 2020
30. http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf, accessed on April, 21st, 2020
31. https://www.kaggle.com, accessed on May, 8th, 2020
32. https://en.unesco.org/covid19/educationresponse/globalcoalition, June, 1st, 2020
33. https://www.statista.com/statistics/982135/smartphone-connections-by-country, accessed on May, 10th, 2020
34. https://www.worldbank.org/en/topic/edutech/brief/how-countries-are-using-edtech-to-support-remote-learning-during-the-covid-19-pandemic, accessed on May, 10th, 2020
35. https://courses.lumenlearning.com/suny-educationalpsychology/chapter/cognitive-development-the-theory-of-jean-piaget, accessed on April, 20th, 2020
36. https://www.apa.org/ed/schools/teaching-learning/top-twenty/principles/assessment, accessed on April, 20th, 2020
37. https://mhrd.gov.in/sites/upload_files/mhrd/files/statistics-new/ESAG-2018.pdf, accessed on May, 15th, 2020
38. http://www.mospi.gov.in/national-statistical-office-nso, accessed on May, 15th, 2020
39. http://www.dqindia.com/powered-by-govt-it-spending-india-to-become-second-largest-it-marketin-apac-region-by-2018, accessed on May, 15th, 2020
40. An intelligent approach to predict the student behaviour and performance. Engineering Technology. 2017; 8(04):5.