Chapter

Information and Communication-Based Collaborative Learning and Behavior Modeling Using Machine Learning Algorithm

Nityashree Nadar and R. Kamatchi

Abstract

Rapid growth of smart phone industries has led people to use more technology and thus aided in adoption of information and communication technology (ICT) in educational purposes for enhancing students' performance. This chapter shows that students use social media platform or virtual environment for learning, especially in Open University or online learning system. In such environment, the students' drop rate is extremely high. This work primarily aims at reducing students' dropout or students' fails to finish course within prerequisite time using student behavior styles. For addressing research problems, this research aims in building efficient student behavior learning model for improving the performance of student applying machine learning (ML) models. The behavior extraction and study have been carried utilizing decision tree (DT) ML algorithm. Further, a model has been proposed for provisioning student contextual information to different students utilizing VLE platform interaction (collaborative learning) using DT algorithm which considered bagging. The DT with bagging is an ensemble learning (EL) model that depicts bootstrap aggregating (BA), which is modeled for enhancing accuracies and stabilities of every distinct predictive trees. Bagging aids DT in influencing overfitting problems and minimizes its variance. The proposed method is efficient in extracting learning styles and intrinsic behavior of students.

Keywords: behavior modeling, information and communication technology, machine learning, student learning style, virtual learning environment

1. Introduction

This chapter presents collaborative learning model to extract behavior and learning style of students. This chapter describes set of learning style for extracting behavior of students. Further, it also discusses how collaborative learning model aids in designing or understanding behavior of student so as to optimize its training program. Further, this chapter shows how using machine learning aid in increasing accuracy of behavior classification. Along with, presents a student learning style intrinsic behavior classification model using decision tree algorithm with bagging. Mathematical model of proposed decision tree algorithm and decision tree algorithm with bagging is given. Then, the experimental evaluation and result attained by proposed model over existing model is described. Lastly, the overall summary of the chapter is given.

The main objective of this study is to recognize different sorts of ICT teaching methods established in market and determine the role, efficiency and competencies of these practices. The enhancing the ICT practices are characterized as initiatives, activities or projects that have tangible impact on teaching skills and humanizing between learner and trainer. For building an efficient learning model in enhancing performance (i.e., reducing drop rate) of students in academics using ICT [48–50].

Firstly, many analysts explained that learning styles promote that increase in knowledge, and make knowing smoother concerning towards students. Learning handling processed include quite prospering as part of e-learning equal, though a thing executed maybe not incorporate learning trends. Since learning styles should feel secure inside thinking of handling processes, Students 'behavior at the internet Program Requires feel, which is examined or followed. In one of the outset in this work, various programs in which students among various learning models operate in a different way additionally strategy suggestions and different approaches are defined. Second one, emerging approach for the learning styles is proposed by employing machine learning technique. **Figure 1**, depicts concerts Platform Goals and **Figure 2**, the proposed envisioned design is shown. It shows a personalized content delivery of e-learning or online-based learning environment that combine data from student, social media, instructor, classification model [1–3], and content to provision personalized course [6–8].

Secondly, social media has been considered and generally defined as the medium through which information is transmitted among various learners and research communities. This social media platform has been used by various educational institutions for encouraging students to collaboratively learn and interact socially. This work studies and examines the usage of social media in the process of collaborative-based learning using learning algebraic math. In this work, different factors considered to enhancing collaborative learning to study algebra of context

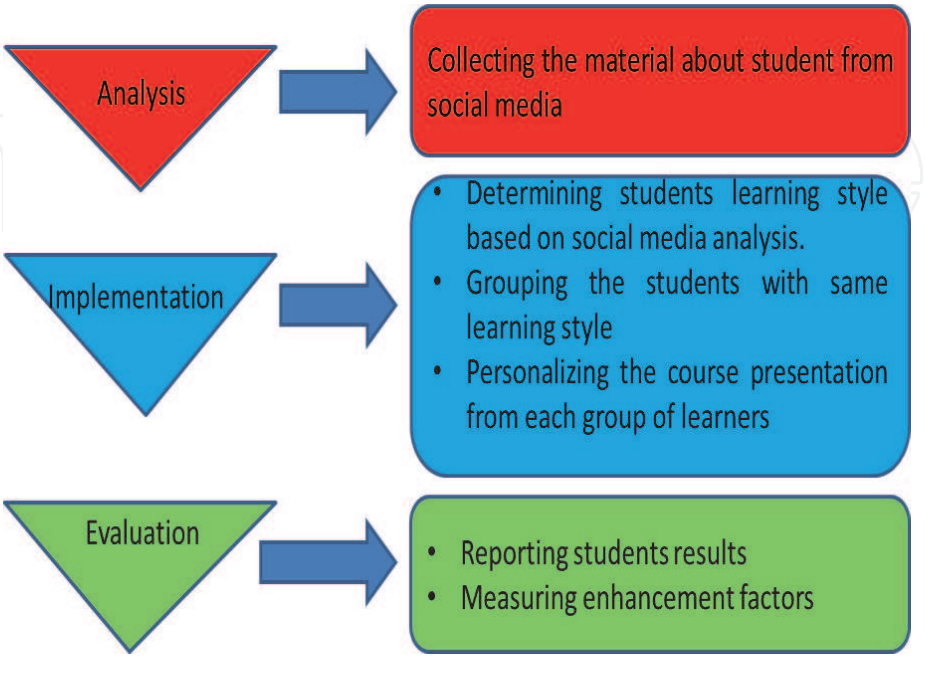


Figure 1.
Proposed framework milestones.

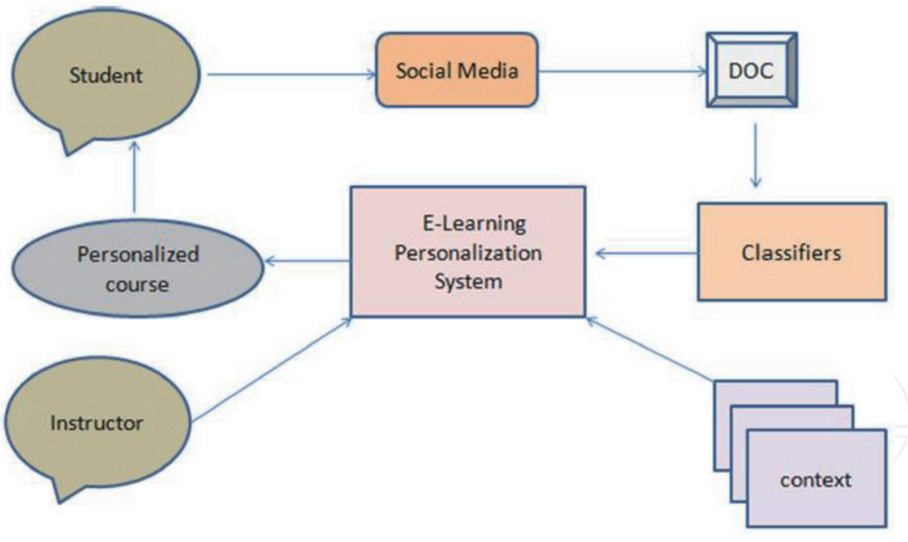


Figure 2.
EMDL (E-learning model design for online learning systems).

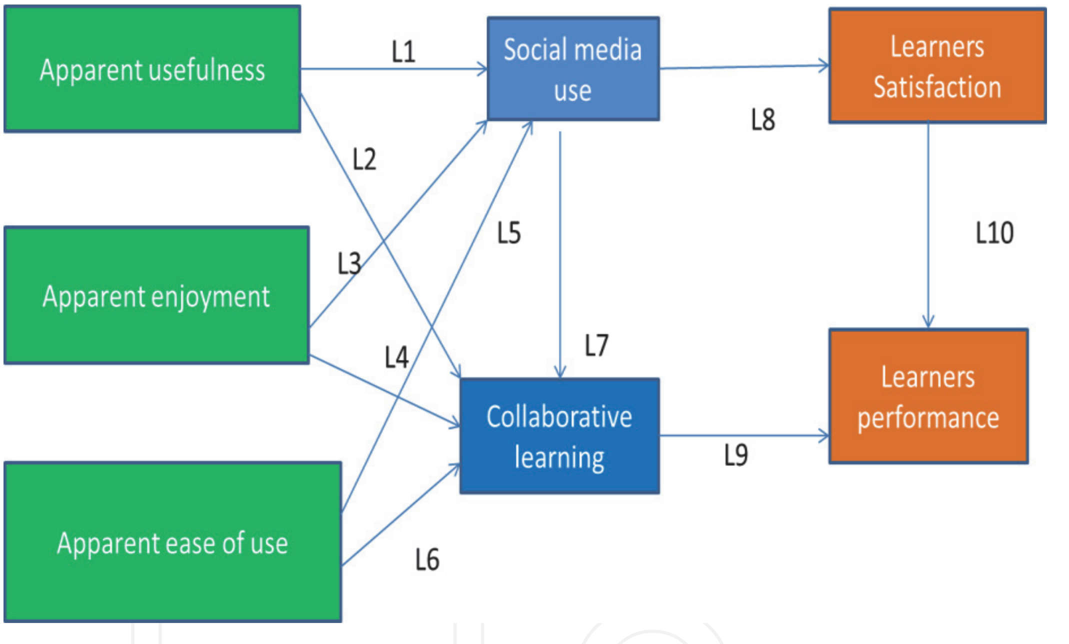


Figure 3.
Proposed hybrid architecture.

using social media are going to be examined. The chapter presents the hybrid research model for ICT education using social media (virtual learning environment) and collaborative learning as shown in **Figure 3**, and the proposed model is analyzed by using machine learning techniques.

2. Proposed collaborative learning using machine learning technique

This section presents proposed collaborative learning model using machine learning technique. Firstly, the behavior and learning style of student are extracted. Then, collaborative learning mechanism using machine learning algorithm using decision tree is presented to accurately extract behavior of student. Lastly, intrinsic behavior of student is extracted using decision tree with bagging method.

2.1 Behavior and earning styles

Regarding learning, this work identified that not all student or person learns in similar manner. A particular set of learning abilities is possessed by each person; thus, the preferences can be identified that constitute his or her learning style. "One size do not fit all" is convey to us by educational research. The learning characteristics of all students differ as informed to us by educational research [9]. Educational research suggests that student's process learns and represents knowledge in different ways, and they prefer to use different type of resources. It is also suggested by the research that it is possible to diagnose students learning style. Further, when instruction is referred to the way they learn some students learn more effectively [10]. Both teacher and students are aided by knowing their learning styles. Teaching and learning strategies can be elaborated better in order to allow students to assimilate in an effective manner and more efficient way in gaining new knowledge and information. To identify and implement better teaching and learning strategies understanding of learning styles can be used [11, 12].

The aforementioned description is a representative of serious mismatches among the teaching style of the instructor and the learning style of student. Students tend to get bored and inattentive, perform poorly on tests, and get discouraged about the course when such a mismatch occurs. This may conclude the students thinking as to withdraw in the subject or course [13]. Some researches in the area of learning styles advocate teaching and learning styles to be matched and bridging the gap between teacher and learners perception in order to reduce teacher's student's style conflicts [14, 15]. This plays an important role enabling students to maximize their classroom experience.

2.2 Learning styles

The theory of learning style depicts the fact that each person has his or her own method or set of Strategies for learning. Preferences strength and characteristic in the way people receive and process information is the definition of a learning style. The strategy may differ from one person to another but it has been narrowed down only up to the GT (Global Trends). The particular ways of learning and this Global trends constitutes the learning style [16]. The trends that all individual have different learning style can be established in a classroom.

Similar lesson has been given to the other student groups, few of them has better performances assured than others students. According to [17], there are several theories about learning styles [17]. Students are classified by a model of learning styles according to a scale that reflects the way the process and receive information. Different learning styles classifies the students, this is done in accordance with the scaling which reflects on similar way as they receive and process the information. Where there is many number of LS tools as well as methodologies [12], two similar assessment equipment's are said to be predominant in the science as well as engineering education Kolb's Learning Styles Inventory (LSI) [18] and the Soloman–Felder Index of Learning Styles (ILS) [11]. For the base of study the fielder and Solomon model was chosen on the study basis of [16] because the other is approved it and it has been implemented in the other work as well [19–21]. Moreover the other researcher [22, 23], because results are easy to interpret and it is user friendly and because the number of dimensions is controlled and can actually be implemented [21].

2.3 Types of behavior and learning style used

The work consider various kinds of learning styles such as active, sensitive, intuitive, visual, global, verbal, reflective, and sequential [5]. Sensitive learning style: Here the courses should poses direct connection with real-time or actual world application content. Intuitive learning style: Here the study material should be designed in theoretical manner with meaning. Along with, it should be innovative with mathematical formula, with proper abstraction and no repetition of content. Visual learning style: here the study material should have lot of visual (i.e., figures and blocks) that depicts certain action. The visualization aid student to remember, understand the concept more easily. Verbal learning style: Here the study material should possess lot of oral presentation with textual data. This kind of student can be given a small abstract to describe or summarize it. Active learning style: Here the learner tends to learn new concept and integrate with practice through discussion. These kind of student should be given assignment in a collaborative manner. Reflexive learning style: This Learning style is based on students' observation and experiences, collection and analyzation of data is done. Prior to making any decision study material must be related with the experiences, personal work must also be included in the requested homework. Sequential learning style: Here the contents are given in steps and chapter wise. The steps or chapter must be logically divided and well connected. Global learning style: Here the assignments are given in random manner. This makes student to think in innovative manner and solve problems in quick manner but may have difficulties to explain how they did it. The above learning styles play an important part in improving student performance.

2.4 Individual and collaborative learning model

Learning style (LS) is characterized as trademark qualities and inclinations in the manners individuals understand and process data [24]. Every student has his/her own method for learning and understanding. However, the most existing model [48–50] is designed for forecasting student drop rate in school. Thus, these model cannot be used online and collaborative environment. In virtual learning environment, student can be separately assisted and their particular requirement can be satisfied through learning process. So as to do that, it is fundamental for VLE platform to keep the data about the students that is viewed as significant for adaptive learning procedure in the student learning environment. Among the student behavior feature sets, used in student learning models, learning styles (LS) comprise a significant environment for improving student specific learning [25–27]. Students gain knowledge from their distinct cooperation/interaction with learning contents; however, they can likewise get information while completing task in collaborative manner with other students. Grouping different individual student based on their characteristic, behavior feature sets using leaning styles can aid in proving the outcome of student learning process [28, 29].

The adaptive learning based on student specific (i.e., based on contextual behavior) is carried out using students learning style, is: (i) sequential students must be straightforwardly assisted through study material, since global students ought to have the option to examine the course in a global manner prior to contemplating certain subjects and (ii) sensing learners will in general like to watch and associate with models prior to learning mathematical ideas or strategies, while instinctive students generally desire the opposite manner. The autonomous grouping is completed in two stages [30]. First, grouping rules decide the clustering arrangement with respect to the individual feature sets and inclinations of the learners. Second, for every collaborative assignment, when it is accessible to a base number of people belongs with similar group, sub-group sets are created and student can start the cooperation/collaboration activity. Similar to many state-of-art model, here we consider the distance metric among individual group member as a significant element for deciding the grouping rules. The final distance is obtained by summing the Euclidean distance (ED), the sensing-intuitive distance and active-reflective distance. For instance, if learner x has gotten the *ILS* score $(x_1 \dots x_2 \dots x_3 \dots x_4)$ and learner y $(y_1 \dots y_2 \dots y_3 \dots y_4)$, the distance D between them is:

$$D = EuclDist + ActRefDist + SenIntDist \tag{1}$$

The Euclidian distance is computed using following equation

$$EuclDist = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2} \tag{2}$$

Then, the active with respect to reflective distance is computed using following equation

$$ActRefDist = \sqrt{(x_1 - y_1)^2} \tag{3}$$

Similarly, the sensitive with respect to intuitive distance is computed using following equation

$$SenIntDist = \sqrt{(x_1 - y_1)^2} \tag{4}$$

Results, interpreted in many state-of-art models which suggests that the pairs of learners with distance matric greater than the average gets superior outcome results than pairs lower than that. Considering the case, the process for choosing collaborators is to choose arbitrarily the preliminary associate and the at last, compute the farthermost conceivable collaborator to the members of a particular group. However, the learning efficiency can be improved by using machine learning model.

2.5 Using machine learning for obtaining enhanced collaborative learning model

One of the major features that play a critical part in establishing knowledge is learning style. This depicts to the respective student way in which they model a learning assignment. The relation between student behavior and learning styles as online VLE environment or other old-fashioned education frameworks has been investigated by many studies. It is motivating to discover the importance of behavior characteristics of student's practice pattern of this virtual learning environment.

Web-based learning management systems are extensively used nowadays and produce vast amounts of data that are potentially useful for improving educational process [31, 32]. The new evolving area, called Educational Data Mining (EDM), concerns with modeling designs that determine knowledge from information coming from various other sources (traditional or distance learning) environments [33]. Increasing research interests in utilizing DM in teaching is recorded in the last decade [34, 35, 36] with focus on various features of teaching procedures (e.g. teacher, student, study material, etc.). Benefits from extracting knowledge from e-learning data are expected under assumption that the trails of user actions can be used to identify specific information on users. We hope that the user behavior captured in log files and recorded in data structures can be used to create models that predict user behavior, or describe their peculiarities. There are several groups of people who can leverage this knowledge, and are potential stakeholders: Students, Teachers, E-learning system administrators and University management. These stakeholders could use this knowledge for different goals [35]:

- Application that mainly deals with student assessment learning performance.
- Application which gives the Learning recommendation (LR) and Course adoption (CA), these are based on the Student Learning Behavior (SLB).
- Various approaches which deals with learning material evaluation and other educational courses based on the online virtual learning environment.
- Other application which involves the straight feedback for the both students and teachers of e-learning courses; these courses were based on the behavior of student learning.
- Developing the detection nature of a typical student's behavior of learning.

These goals are achieved with the help of data mining techniques such as support vector machines, artificial neural networks, decision trees, k-nearest neighbor, hierarchical clustering, K-means etc. [37]. Since data is not stored in a systematic way as learning management systems (LMS) are not primarily designed with data analysis and Mining in mind. For performing thorough analysis on such data, long and tedious pre-processing is required. Statistical reports are usually produced by LMS systems. Useful conclusions cannot be drawn from this instructions idea for the course potential for student abilities and they are useful only for platform administrative purposes. This chapter describes how one can influence the available data on student learning style and student behavior, in order to forecast success of students, as well as profile students into clusters which may aid in enhancing state-of-art learning content, method, and collaborative learning (CL).

2.6 Using decision tree machine learning algorithm for obtaining enhanced collaborative learning model

The decision tree (DT) is a data mining method or strategy for tackling grouping, classifying and forecasting issues. DT are a basic iterative/recursive building block for communicating a successive classifying operation in which a case, portrayed by a lot of attributes sets, is appointed to one of a disjoint features of classes. DT comprise of nodes (parent) and leaves (child/sibling). Every parent in the tree includes testing a specific characteristic and each leaf of the tree means a class. More often than not, the test contrasts an attributes with a constant. Child nodes give an order that applies to all occurrences that achieve the child, or a classification sets, or a likelihood dispersion over every conceivable clustering. For carrying out classification operation on obscure occasion or condition, it is traversed down the tree as per the estimations of the attributes feature tried in progressive nodes, and when a leaf is achieved, the case is arranged by the class allocated to the leaf. In the event that the property that is tried at a node is an ostensible one, the quantity of siblings is generally the quantity of conceivable estimations of the feature set. The tree intricacy is estimated by one of the accompanying measurements: the all-out number of nodes, complete number of leaves, tree size level and number of feature set utilized [38–41].

As referenced previously, the issue of developing a DT can be communicated iteratively. In the first place, it is important to choose a feature set to put at the root node, and make one branch for every conceivable parameter. This parts up the model set into subsets, one for each estimation of the feature set. Presently the procedure can be rehashed recursively for each branch, utilizing just those cases that really achieve the branch. In the event that whenever all examples at a node have a similar classifying outcome, that piece of the tree needs to end creating [39]. As indicated by [41, 42], the way establishing the feature set that delivers the best split in the information is the one of the primary contrasts among the different DT construction methods.

There are a few method of segmenting/splitting measures. Every DT method utilizes its very own measure to choose among the feature set at each progression while developing or constructing the tree. This work use J48 which is a usage of C4.5 calculation was developed in 1992, by Ross Quinlan, to beat the confinement of the ID3 method (inaccessible qualities, continuous feature set estimation ranges, pruning of DT, and so on.) [43]. C4.5 utilizes a divide-and-conquer way to deal with developing DT. The gain ratio is the default segmenting rule utilized by C4.5, a data-based measure that considers diverse number of test results [44].

$$GR(S, A) = \frac{G(K, H)}{Split\ Info\ (K, H)} \tag{5}$$

where parameter (H) is the set of all conceivable or probable parameter for feature sets H, and K_w is the subset of K for which feature set H has parameter w.

As indicated by [38], the focal decision any tree-based method is choosing which feature set to test at every node in the tree. There is a decent quantifiable metric for this issue, called data or information gain (IG). Yet, so as to characterize IG decisively, it is important to characterize a measure usually utilized in data hypothesis, depicted as entropy, that describes the (im)purity of a self-assertive or random gathering (sets) of models. In the event that the objective characteristic can take on m distinctive parameter, at that point the entropy of K in respect to this m-wise classifying outcome is characterized as [38]:

$$Entropy\ (K) = -\sum_{j=1}^{m} p_j \log_2 p_j \tag{6}$$

where K is an assumed collection set, and p_j is the ratio of K belongs to class label j.

The assumed entropy as a proportion of the (im)purity influence in an accumulation of training models, a proportion of the viability of a feature set in performing classification on training information can be characterized now. The measure is called IG. It is the normal decrease in entropy brought about by apportioning the models as per this quality (feature sets). The IG, $G(K, H)$ of a feature set H, comparing to a dataset of precedents K, is depicted as pursues

$$G(K, H) = Entropy\ (K) - \sum_{w \in parameter(H)} \frac{|K_w|}{K} Entropy\ (K_w) \tag{7}$$

Using proposed decision tree algorithm (machine learning) enhances the collaborative learning efficiency of student behavior and learning styles which is experimentally shown n later section.

2.7 Intrinsic behavior model using proposed decision tree-based classification model

For extracting intrinsic behavior of student, it is important to identify correlation factor among learning styles of students. Further, using correlation measure this work uses decision tree algorithm with bagging to extract intrinsic behavior of student learning style. This work uses decision tree classification algorithm in previous section, however with small optimization introduced. The proposed classification algorithm is an ensemble machine learning (EML) methodology that works by building a large number of DT at training instance and resulting in a classification outcome, which is averaged by each distinct tree [45]. This work uses an additional information or layer of arbitrariness to bagging technique similar to method presented by [47]. Our method will not only performs well for classification and regression function, and at the same time it efficient in identifying behavior [46].

Bagging, which alludes to bootstrap aggregation function, is an ELM method intended to improve the dependability and precision of distinct prescient or forecasting method, for example, trees [47]. Bagging encourages DT to diminish their fluctuation and the impact of overfitting. Considering that a preparation/training dataset is given by $A = a_{1,2\,\ldots\,n}$ with response $vB = a_{1,2\,\ldots\,n}$, bagging will repeat L instance to choose a self-assertive or random example with substitution of the preparation dataset and fits trees to these feature set sample. A tree i_l, $(L = 1, 2, \ldots L)$ will be trained each instance. Subsequent to preparing, the obtained forecasting method model can be built up by meaning the forecasts from L regression trees (RT) or by taking the greater part vote from L DT. Note that examples are chosen with substitution, and the likelihood that a specific feature set sample isn't chosen after K instance determination can be depicted as follows

$$P = \left(1 - \frac{1}{n}\right)^{L} \tag{8}$$

In the bagging procedure of proposed grouping (classifying) method, L as a rule equivalents to n. At the point when n is sufficiently enormous, and they are gotten out-of-bags feature set tests.

Furthermore, proposed classification method enhance the universal tree developing method, where at every applicant split in the tree method, an arbitrary subset of the highlights are utilized as opposed to choosing a specific component from every one of the hopefuls. Though in a conventional tree ELM method, if a couple of attributes are solid indicators for the forecasting, these attributes sets will be chosen in a large number of the base estimators. At that point, these trees will be greatly corresponded, hence debilitating the forecast capacities.

The hypothetical foundation of proposed classifying model can be essentially isolated into two sections: RF convergence hypothesis and generalization error limit (GEL). All the evidence of strategy can be obtained from [45]. Subsequently, the convergence of GLE demonstrates that random forest can deliver a restricting estimation of the GE and do not overfit as more trees are included. The upper limit for the GE is obtained using following equation

$$GE \leq \frac{\beta(1 - t^2)}{t^2} \tag{9}$$

where β is the average estimation of the relationship, and t is the quality of a distinct tree in the RF display. It implies that with expanding the quality of distinct tree and lessening the relationship between trees, the proposed classifying method will accomplish progressively exact forecast results.

In addition, as portrayed above, so as to expand the distinct tree quality in the proposed classifying method, attributes set investigation must be first done to decide the overwhelming conduct/learning style for each errand. As it were, good attribute selection, conduct estimations or mix of learning style, firmly identified with every particular conduct must be planned before positioning the attributes sets. At that point, in view of the out-of-bag test sets, every one of the attributes sets can be arranged by the forecast capacity with the out of-bag computes. All the more explicitly, tree-organized classification model in proposed classifier method that have essential factors at nodes should be exceptionally identified with the reaction, so imperative factors can be chosen in these solid trees.

3. Result and analysis

This section evaluates performance evaluation of proposed student behavior (learning style) learning model over existing models. To carry out the experimental analysis, similar to case study (dataset) [5] is considered. However, the dataset is not publically available. As a result, we have generated a dataset similar to [5]. The research work aimed at studying different actions in group (i.e., with different VLE platform). Even though there are different kind of actions in VLE environment, this work have used a total number of actions that majority of learner will carry out on varied VLE platforms. The work studied at predicting behavior and leaning styles of a student that a learner will utilize for collaborating, communicating, and learning assistance. For experiment analysis we have collected data from student considering questionaries' (task) such as presentation, topic discussion etc. More details of dataset can be obtained from Section 3.1 and 3.2. The experiment is conducted using windows 10 OS, 3.2 GHz Intel quad core processor with 16 GB RAM. For extracting behavior, using these collected data ML classifier, using decision tree algorithm is applied. This work use Precision, Recall, ROC, and F-measure to evaluate performance. The precision P^r is computed as follows

$$P^r = \frac{T^p}{T^p + F^p} \tag{10}$$

where T^p is true positive and F^p is false positive.

The recall R^c is computed as follows

$$P^r = \frac{T^p}{T^p + F^n} \tag{11}$$

where F^n is false negative.

The F-measure F is calculated as follows

$$F = \frac{2 * P^r * R^c}{P^r + R^c} \tag{12}$$

Similarly, the ROC is computed as follows

$$ROC = \frac{T^p + T^n}{T^p + T^n + F^p + F^n} \tag{13}$$

where T^n is true negative and F^n is false negative.

The specificity S is calculated as follows

$$S = \frac{T^n}{T^n + F^p} \tag{14}$$

3.1 Problem, issues, and challenges faced in collecting student data

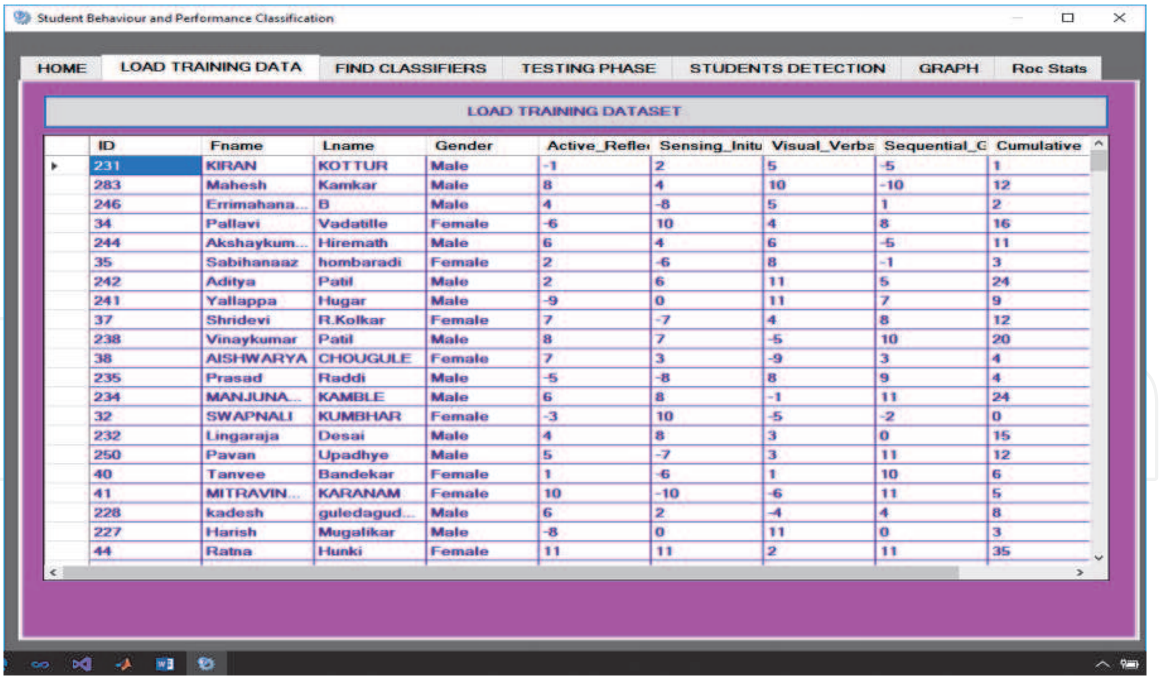
Collecting open source ICT data pertaining to student is challenging. Initially we found difficulties/issues in collecting ICT data of student as there was no any publically available open source dataset. Thus, we have conducted extensive survey of various existing methodology. Form the survey it is evident that many existing methods have manually generated data for extracting learning style of students. Using these styles, the behavior pattern of students is extracted, and the teaching styles are optimized. Further, few model have used social circle for collaborative learning to enhance student behavior. Collecting these kind of data was very much challenging. As a result, we have generated our data for behavior modeling, considering various tasks and assessments. We have collected social circle of various user in Facebook and Twitter and performed collaborative analysis on them. However, it is of very limited use for adopting the student learning performance. Finally, after extensive searching we were able to find an open source data where student online interactions (virtual learning environment) are collected. These data are collected and stored in local hard disk. However, these data have to be anonymized and proper preprocessing steps must be done to support usage of machine learning algorithm on them.

3.2 Data collection for student behavior modeling for conducting experiment analysis

For analysis student learning skills, we have collected student's data, which are composed of data such as number of actions performed for each category (that is, considering different social media platform). The number of actions in this research work has only considered the total number of actions which a student performs on different platforms. Data collection is done considering examining whether it is likely to forecast a student's learning style using their behavior that a student utilizes for communication, collaboration and learning support. Initially, the data is collected considering three kinds of behaviors such as verbal, visual and global. Later, we have collected considering other behavior such as active, reflective, sequential, sensing, and intuitive. Descriptions about each behavior are given in Sections 2.1 and 2.2. The data is collected from students considering questionaries' (task) such as presentation, topic discussion etc. Each student must complete at least one task. Then, the teacher makes assessment of these tasks and gives score that ranges −15 to −15 for different behaviors. Total dataset is composed of 400 student. The sample of data collected is described in **Figure 4**. Further, **Figures 5** and **6** describes the labeled dataset which composed of two classes A and B. Class A depicts the class with positive performance (i.e., score greater than zero) and Class B depicts the class with negative performance (i.e., score lesser than zero). Thus, identifying what type of class aid in improving student learning performance.

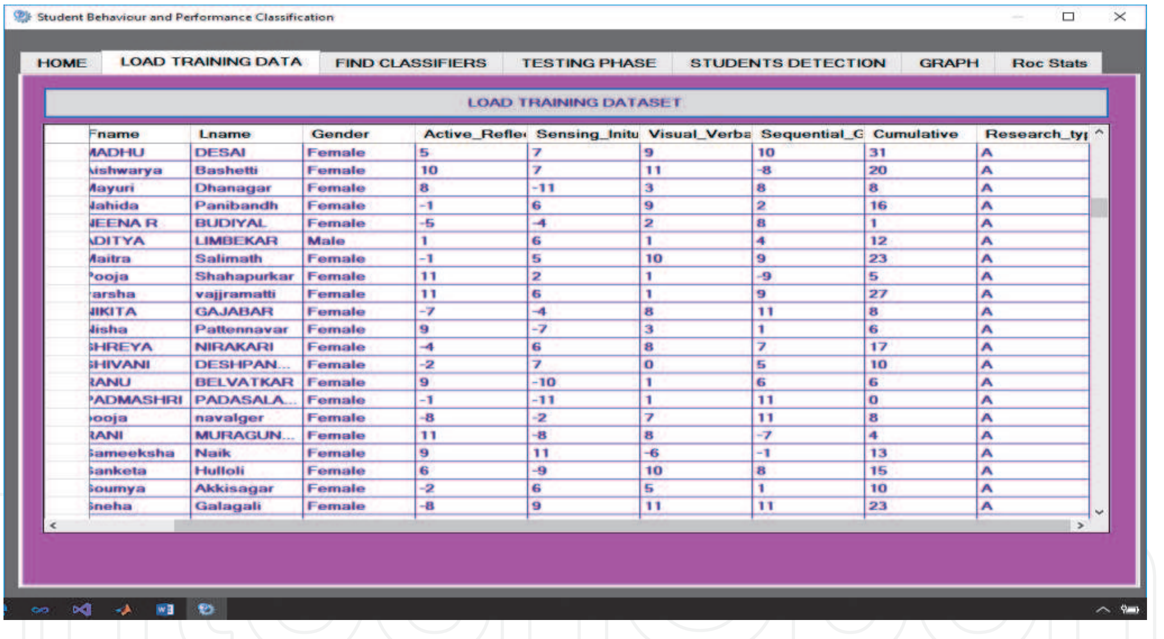
3.3 Precision performance evaluation

This section carries out experiment analysis for evaluating performance of proposed classification method using ML over existing ML-based classification method in terms of precision. The **Figure 7**, displays precision result accomplished by



ID	Fname	Lname	Gender	Active_Refle	Sensing_Initu	Visual_Verba	Sequential_G	Cumulative
231	KIRAN	KOTTUR	Male	-1	2	5	-5	1
283	Mahesh	Kamkar	Male	8	4	10	-10	12
246	Errimahana...	B	Male	4	-8	5	1	2
34	Pallavi	Vadatille	Female	-6	10	4	8	16
244	Akshaykum...	Hiremath	Male	6	4	6	-5	11
35	Sabihanaaz	hombaradi	Female	2	-6	8	-1	3
242	Aditya	Patil	Male	2	6	11	5	24
241	Yallappa	Hugar	Male	-9	0	11	7	9
37	Shridevi	R.Kolkar	Female	7	-7	4	8	12
238	Vinaykumar	Patil	Male	8	7	-5	10	20
38	AISHWARYA	CHOUGULE	Female	7	3	-9	3	4
235	Prasad	Raddi	Male	-5	-8	8	9	4
234	MANJUNA...	KAMBLE	Male	6	8	-1	11	24
32	SWAPNALI	KUMBHAR	Female	-3	10	-5	-2	0
232	Lingaraja	Desai	Male	4	8	3	0	15
250	Pavan	Upadhye	Male	5	-7	3	11	12
40	Tanvee	Bandekar	Female	1	-6	1	10	6
41	MITRAVIN...	KARANAM	Female	10	-10	-6	11	5
228	kadesh	guledagud...	Male	6	2	-4	4	8
227	Harish	Mugalikar	Male	-8	0	11	0	3
44	Ratna	Hunki	Female	11	11	2	11	35

Figure 4.
The sample of data collected considering different behaviors of both female and male.



Fname	Lname	Gender	Active_Refle	Sensing_Initu	Visual_Verba	Sequential_G	Cumulative	Research_typ
ADHU	DESAI	Female	5	7	9	10	31	A
ishwarya	Bashetti	Female	10	7	11	-8	20	A
ayuri	Dhanagar	Female	8	-11	3	8	8	A
ahida	Panibandh	Female	-1	6	9	2	16	A
EENA R	BUDIYAL	Female	-5	-4	2	8	1	A
DITYA	LIMBEKAR	Male	1	6	1	4	12	A
aitra	Salimath	Female	-1	5	10	9	23	A
ooja	Shahapurkar	Female	11	2	1	-9	5	A
arsha	vajjramatti	Female	11	6	1	9	27	A
IKITA	GAJABAR	Female	-7	-4	8	11	8	A
isha	Pattennavar	Female	9	-7	3	1	6	A
HREYA	NIRAKARI	Female	-4	6	8	7	17	A
HIVANI	DESHPAN...	Female	-2	7	0	5	10	A
ANU	BELVATKAR	Female	9	-10	1	6	6	A
ADMASHRI	PADASALA...	Female	-1	-11	1	11	0	A
ooja	navalger	Female	-8	-2	7	11	8	A
ANI	MURAGUN...	Female	11	-8	8	-7	4	A
ameeksha	Naik	Female	9	11	-6	-1	13	A
anketa	Hulloli	Female	6	-9	10	8	15	A
oumya	Akkisagar	Female	-2	6	5	1	10	A
neha	Galagali	Female	-8	9	11	11	23	A

Figure 5.
The sample of data collected considering different behaviors of both female and male with binary labeled class (A).

proposed classification method over existing classification method. An average precision performance improvement of 2.07% is attained by proposed classification method over existing classification method. From the results attained, it can be seen that proposed classification method accomplishes superior precision result improvement when compared to existing classification method.

3.4 Recall performance evaluation

This section carries out experiment analysis for evaluating performance of proposed classification method using ML over existing ML-based classification method in terms of recall. The **Figure 8**, shows performance outcome attained by proposed

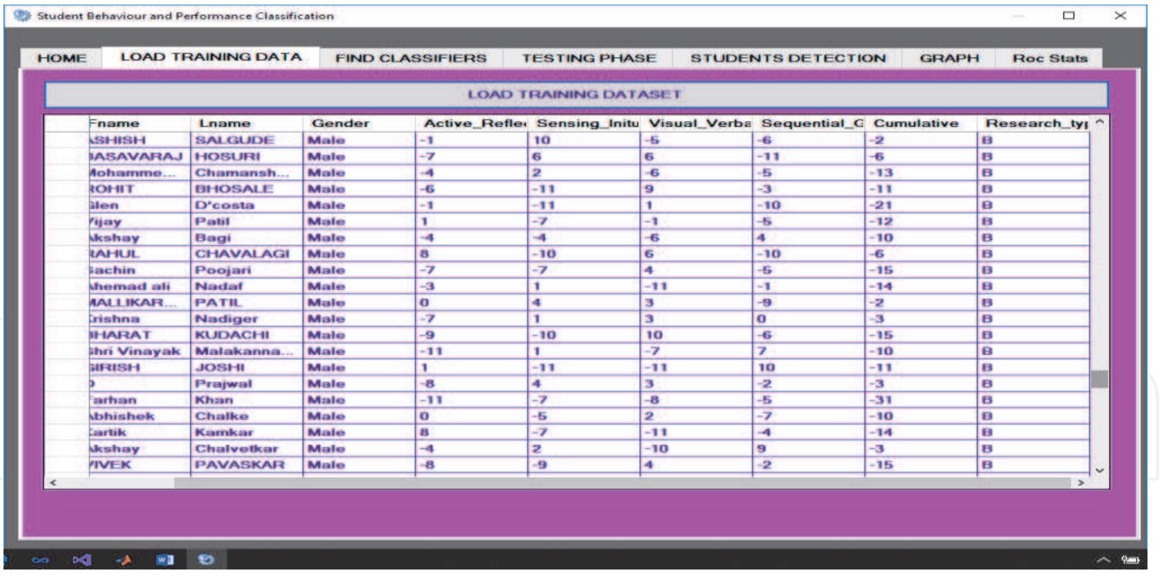


Figure 6.
The sample of data collected considering different behaviors of both female and male with binary labeled class (B).

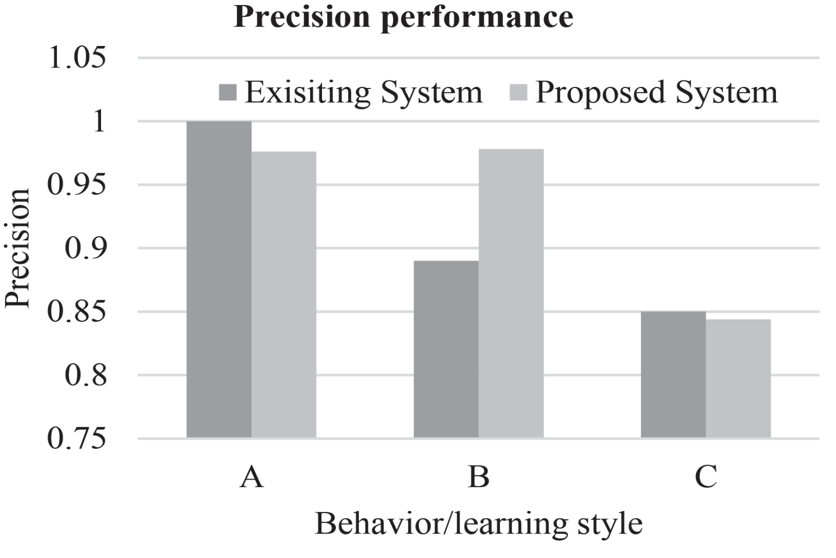


Figure 7.
Precision performance evaluation considering different behavior/learning styles.

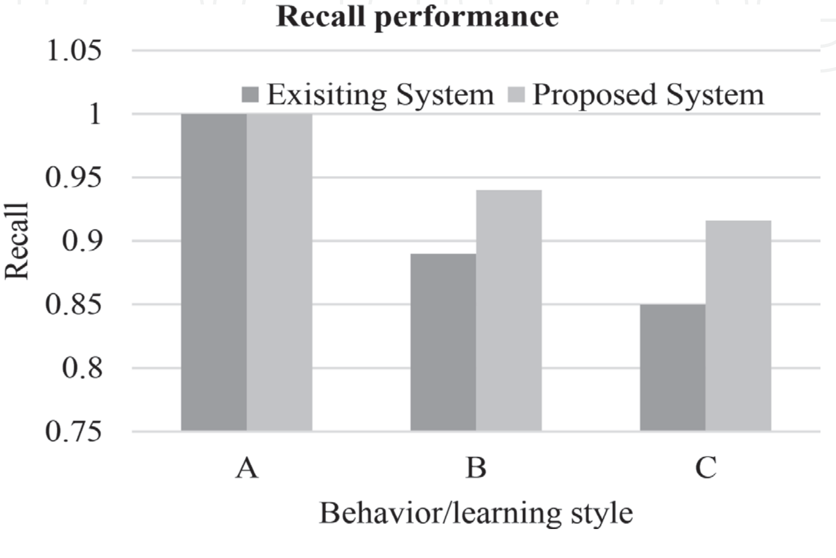


Figure 8.
Recall performance evaluation considering different behavior/learning styles.

classification method and existing classification method in terms of recall. An average recall performance improvement of 3.01% is accomplished by proposed classification method with respect existing classification method. From the results attained, it can be seen, the proposed classification method accomplishes superior recall result improvement than existing classification method.

3.5 F-measure performance evaluation

This section carries out experiment analysis for evaluating performance of proposed classification method using ML over existing ML-based classification method in terms of F-measure. The **Figure 9**, shows F-measure result accomplished by proposed classification method and existing classification method. An average F-measure performance enhancement of 1.53% is accomplished by proposed classification method model with respect to existing classification method. From the results attained, it can be seen the proposed classification method accomplishes superior F-measure result with respect to existing classification method.

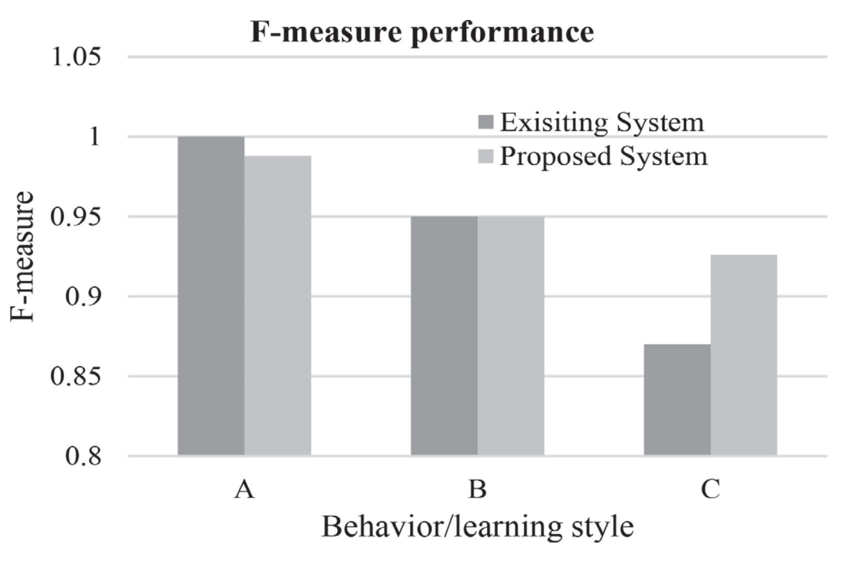


Figure 9.
F-measure performance evaluation considering different behavior/learning styles.

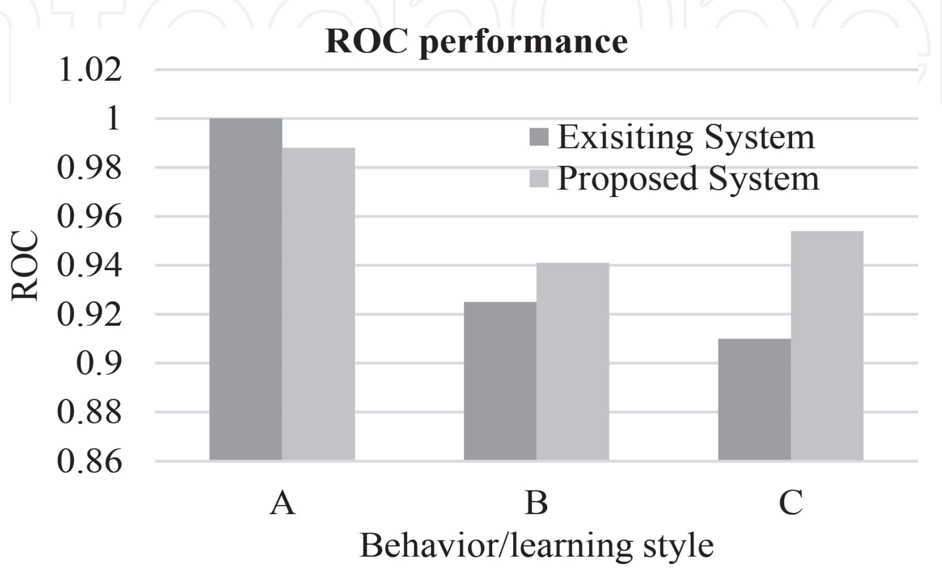


Figure 10.
ROC performance evaluation considering different behavior/learning styles.

3.6 ROC performance evaluation

This section carries out experiment analysis for evaluating performance of proposed classification method using ML over existing ML-based classification method in terms of ROC. The **Figure 10**, displays ROC result accomplished by proposed classification method and existing classification method. An average ROC performance improvement of 1.66% is accomplished by proposed classification method with respect to existing classification method. From the results attained, it can be seen the proposed classification method accomplishes superior ROC results improvement with respect to existing classification method.

3.7 Intrinsic behavior analysis

This section carries out experiment analysis to extract intrinsic behavior of student learning style. We have considered eight behavior style or actions such as

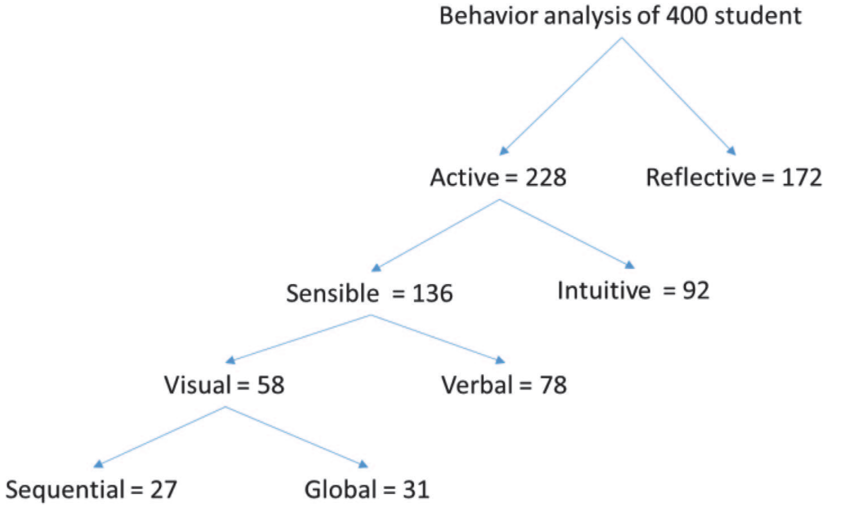


Figure 11.
Decision tree obtained using proposed model to identify how many active students are sensible, visual, sequential and global.

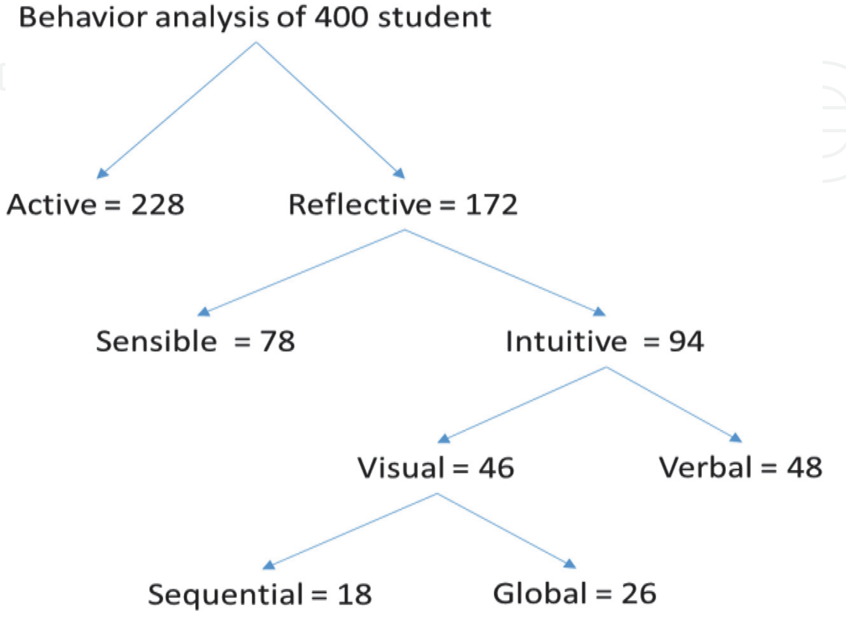


Figure 12.
Decision tree obtained using proposed model to identify how many reflective students are intuitive, visual, sequential and global.

sensing, intuitive, active, reflective, visual, verbal, sequential, and global for experiment analysis. Total dataset composed of 400 students, is used for analysis and experiment is conducted to create decision tree. Various decision tree is built to analyze the behavior and learning style of student as shown in **Figures 11** and **12**. The outcome attained by the proposed model is shown in **Table 1**, and is graphically shown in **Figure 13**. The performance outcome of proposed model over exiting model is shown in T**ables 2–4** considering 100, 200, and 400 students respectively. From the result attained, it can be seen that the proposed model attain good performance. The exiting model did not considered extracting intrinsic behavior of

Reflective	Active	Sensing	Intuitive	Visual	Verbal	Sequential	Global
172	206	196	179	206	183	189	196

Table 1.
Experiment analysis of student learning style.

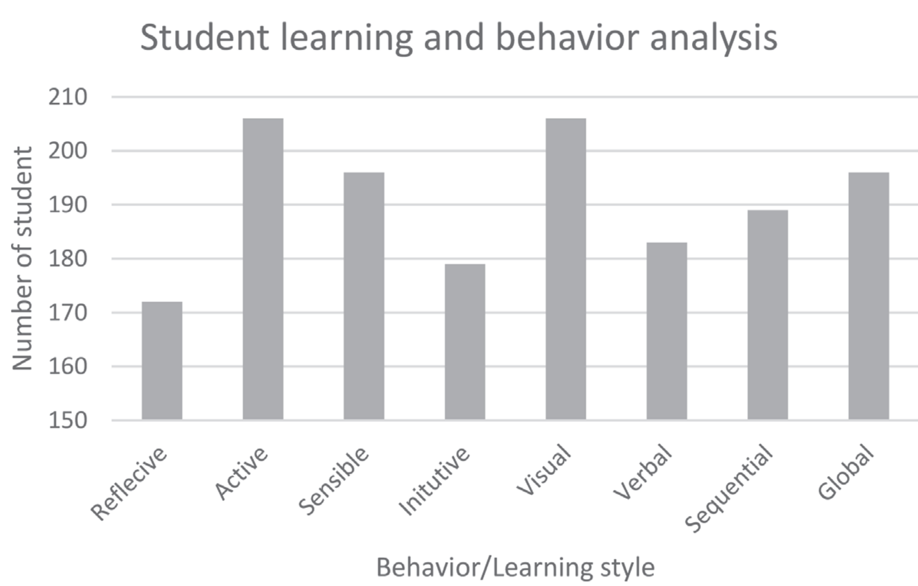


Figure 13.
Student behavior and learning style analysis.

	Accuracy (ROC)	Sensitivity (Recall)	Specificity	Positive prediction (Precision)
Exiting approach [4]	80.6%	79.43%	72.2%	80.17
Our approach	**93.97%**	**91.1%**	**95.3%**	**96.33%**

Table 2.
Algorithm performance evaluation considering 100 students.

	Accuracy (ROC)	Sensitivity (Recall)	Specificity	Positive prediction (Precision)
Exiting approach [4]	81.1%	79.82%	71.96%	81.98
Our approach	**94.17%**	**93.62%**	**96.4%**	**97.32%**

Table 3.
Algorithm performance evaluation considering 200 students.

	Accuracy (ROC)	Sensitivity (Recall)	Specificity	Positive prediction (Precision)
Exiting approach [4]	81.8%	80.24%	71.2%	82.17
Our approach	**95.97%**	**94.6%**	**97.5%**	**98.57%**

Table 4.
Algorithm performance evaluation considering 400 students.

student. However, the proposed model extracts intrinsic behavior of student learning styles. This aid is improving in building better learning model.

4. Conclusions

This chapter firstly discusses the importance, issues, challenges, and problems of using ICT in education. Further, it discusses how collaborative learning using online media enhances or impacts students' learning performance. Then, it also shows that extracting behavior and learning style of student aid enhancing students' performance. The chapter presents an efficient collaborative learning model for enhancing students' performance using machine learning. Experiments are conducted using manually collected dataset. Here we consider only visual, verbal, and global learning styles. The behavior extraction and analysis is done using machine learning such as using decision tree algorithm. Experiment outcome shows that the proposed model improves precision performance by 2.07%, recall performance by 3.01%, F-measure performance by 1.53%, and ROC performance by 1.66% over existing model. The overall results attained shows that the proposed decision tree-based classification model can superiorly extract behaviors and learning styles of students in collaborative manner. Further, the work extracted the intrinsic behavior of student using social media (i.e., online virtual learning environment). This work considers different kinds of behavior or learning style of student such as sensing, intuitive, active, reflective, visual, verbal, sequential, and global for experiment analysis. Experiments are conducted to evaluate the performance of proposed model to extract intrinsic behavior of students. Experiment outcome shows that the proposed model attains Precision (positive predicted value) of 97.4%, Recall (sensitivity) of 93.1%, ROC (Accuracy) of 94.7%, and specificity of 96.4%. The overall result attained shows proposed model attains significant performance over existing model considering extracting intrinsic behavior analysis. Future work would consider building an efficient risk identification model of student fails to complete course on time on Open University and online courses portal.

Author details

Nityashree Nadar[1*] and R. Kamatchi[2*]

1 Information Technology/Computer, S.I.W.S College, Wadala, Mumbai, India

2 ISME school of Management and Research, Mumbai, India

*Address all correspondence to: nityashreenadar@yahoo.co.in and rkamatchiiyer@gmail.com

References

[1] Lakkaraju H, Aguiar E, Shan C, Miller D, Bhanpuri N, Ghani R, et al. A machine learning framework to identify students at risk of adverse academic outcomes. KDD'15 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2015: 1909-1918

[2] Bainbridge J, Melitski J, Zahradnik A, Lauria a E, Jayaprakash SM, Baron J. Using learning analytics to predict At-risk students in online graduate public affairs and administration education. The JPAE Messenger. 2015;**21**(2):247-262

[3] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. CoRR, abs/1603.02754; 2016

[4] Raboca HM, Cărbunărean F. Ict in education - exploratory analysis of students' perceptions regarding Ict impact in the educational process. Managerial Challenges of the Contemporary Society. 2014;**7**(2):59

[5] Velázquez F, Lidia A, Cervantes-Pérez F, Saïd A. A quantitative analysis of student learning styles and teacher teachings strategies in a Mexican higher education institution. Journal of Applied Research and Technology. 2012;**10**: 289-308

[6] Kuzilek MH, Zdrahal Z. Open university learning analytics dataset. In: Data Literacy for Learning Analytics Workshop at LAK16, 26th April 2016. Edinburgh, UK; 2016

[7] Hlosta KJ, Zdenek MZ, et al. Scientific Data. 2017;**4**:170171. DOI: 10.1038/sdata.2017.171, 2017

[8] Reigeluth C. A new paradigm of ISD? Educational Technology & Society. 1996;**36**(3):13-20

[9] Honey P, Mumford A. Using your Learning Styles. Maidenhead: Peter Honey; 1986

[10] Rasmussen K. Hypermedia and learning styles: Can performance be influenced? Journal of Multimedia and Hypermedia. 1998;**7**(4):291-308

[11] Felder R. Soloman B. Learning styles and strategies, North Carolina State University [online]. 1993. Available from: http://www.ncsu.edu/felderpublic/ILSdir/ILS.pdf [Accessed on 28-05-2007]

[12] Coffield F, Moseley D, Hall E, Ecclestone K. Learning styles and pedagogy in post–16 learning: A systematic and critical review, Learning and Skills Research Center Report [online]. 2004. Available from: www.LSRC.ac.uk.

[13] Oxford R, Ehrman M, Lavine R. Style wars: Teacher-student style conflicts in the language classroom. In: Magnan S, editor. Challenges in the 1990's for College Foreign Language Programs. Boston: Heinle and Heinle; 1991

[14] Smith L, Renzulli J. Learning style preference: A practical approach for classroom teachers. Theory Into Practice. 1984;**23**(1):45-50

[15] Charkins R, O'Toole D, Wetzel J. Linking teacher and student learning styles with student achievement and attitudes. The Journal of Economic Education. 1985;**16**:111-120

[16] Felder R, Silverman L. Learning and teaching styles in engineering education. Engineering Education. 1988; **78**:674-681

[17] Sewall T. The measurement of learning style: A critique of four assessment tools. ERIC ED267247; 1986

[18] Kolb D. Experimental Learning: Experience as the Source of Learning and Development. Englewood Cliffs, NJ: Prentice-Hall; 1984

[19] Carver C, Howard R, Lane W. Enhancing student learning through hypermedia courseware and incorporation of student learning styles. IEEE Transactions on Education. 1999; **42**:33-38

[20] Hong H, Kinshuk. Adaptation to student learning styles in web based educational systems. In: Proceedings of World Conference on Educational Multimedia, Hypermedia & Telecommunications EDMEDIA. Lugano, Switzerland; 2004. pp. 21-26

[21] Paredes P, Rodríguez P. Considering sensing intuitive dimension to exposition-exemplification in adaptive sequencing. In: Proceedings of the AH2002 Conference. Malaga, Spain; 2002. pp. 556-559

[22] Zywno M. A Contribution to Validation of Score Meaning for Felder-Soloman's Index of Learning Styles, ASEE Conference. Nashville, Tennessee; 2003

[23] Felder R, Spurlin J. Applications, reliability and validity of the index of learning styles. International Journal of Engineering Education. 2005;**21**(1): 103-112

[24] Felder RM. Matters of style. ASEE Prism. 1996;**6**(4):18-23

[25] Paredes P, Rodriguez P. Considering learning styles in adaptive web-based education. In: 6th World Multiconference on Systemics, Cybernetics and Informatics. Orlando, Florida; 2002. pp. 481-485

[26] Hong H, Kinshuk D. Adaptation to student learning styles in web based educational systems. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications. Lugano, Switzerland; 2004. pp. 491-496

[27] Stash N, Cristea A, De Bra P. Authoring of learning styles in adaptive hypermedia: Problems and solutions. In: World Wide Web Conference. NY, USA; 2004. pp. 114-123

[28] Martin E, Paredes P. Using learning styles for dynamic group formation in adaptive collaborative hypermedia systems. In: Workshop Adaptive Hypermedia and Collaborative Web-Based Systems (AHCW04), Web Engineering. Munich, Germany; 2004. pp. 188-197

[29] Deibel K. Team formation methods for increasing interaction during In-class group work. In: 10th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education. Portugal; 2005. pp. 291-295

[30] Carro RM, Ortigosa A, Martin E, Schlichter J. Dynamic generation of adaptive web-based collaborative courses. In: Groupware: Design, Implementation and Use. Vol. 2806. Berlin, Heildelberg; 2003. pp. 191-198. LNCS

[31] Martin-Blas T, Serano-Fernandez A. The role of new technologies in the learning process: Moodle as a teaching tool in physics. Computers & Education. 2009;**52**:35-44. DOI: 10.1016/j.compedu.2008.06.005

[32] García-Peñalvo FJ, Conde MÁ, Alier M, Casany MJ. Opening learning management systems to personal learning environments. Journal of Universal Computer Science. 2011; **17**(9):1222-1240

[33] Romero C, Ventura S, García E. Data mining in course management systems: Moodle case study and tutorial. Computers in Education. 2008;**51**(1): 368-384

[34] Kumar V. An empirical study of the applications of data mining techniques in higher education. International Journal of Advanced Computer Science and Applications. 2011;**2**(3):80-84

[35] Castro F, Vellido A, Nebot À, Mugica F. Applying data mining techniques to e-learning problems. Evolution of Teaching and Learning Paradigms in Intelligent Environment. 2007;**62**:183-221

[36] Weng C-H. Mining fuzzy specific rare itemsets for education data. Knowledge-Based Systems. 2011;**24**(5): 697-708

[37] Romero C, Ventura S. Educational data mining: A review of the state-of-the-art. IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews. 2011;**40**(6): 601-618

[38] Mitchell TM. Machine Learning. India: McGraw-Hill; 1997

[39] Witten IH, Frank E, Hall MA. Data Mining: Practical Machine Learning Tools and Techniques. Burlington: Morgan Kaufmann Publishers; 2011

[40] Quinlan RJ. Generating production rules from decision trees. IJCAI. 1987;**87**: 304-307

[41] Rokach L, Maimon O. Data Mining with Decision Trees: Theory and Applications. World Scientific; 2014

[42] Vandamme J-P, Meskens N, Superby J-F. Predicting academic performance by data mining methods. Education Economics. 2007;**15**(4): 405-419

[43] Hssina B, Merbouha A, Ezzikouri H, Erritali M. A comparative study of decision tree ID3 and C4.5. International Journal of Advanced Computer Science and Applications. 2014;**4**(2):13-19

[44] Quinlan RJ. Improved use of continuous attributes in C4.5. Journal of Arti. 1996;**4**:77-90

[45] Breiman L. Random forests. Machine Learning. 2001;**45**(1):5-32

[46] Genuer R, Poggi JM, et al. Variable selection using random forests. Pattern Recognition Letters. 2010;**31**(14): 2225-1136

[47] Breiman L. Bagging predictors. Machine Learning. 1996;**24**(2):123-140

[48] Cano A, Leonard JD. Interpretable multiview early warning system adapted to underrepresented student populations. IEEE Transactions on Learning Technologies. 2019;**12**(2): 198-211

[49] Márquez-Vera C, Cano A, Romero C, et al. Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. Applied Intelligence. 2013;**38**:315. DOI: 10.1007/s10489-012-0374-8

[50] Márquez-Vera C, Cano A, Romero C, et al. Early dropout prediction using data mining: A case study with high school students. Expert Systems. 2016;**33**(1):107-124. DOI: 10.1111/exsy.12135/abstract