Asian Journal of Research in Computer Science

Volume 17, Issue 3, Page 15-37, 2024; Article no.AJRCOS.112361
ISSN: 2581-8260

Identifying the Learning Style of Students Using Machine Learning Techniques: An Approach of Felder Silverman Learning Style Model (FSLSM)

Wanniarachchi, WAAM [a] and Premadasa, HKS [b*]

[a] Faculty of Computing, General Sir John Kotelawala Defence University, Sri Lanka.
[b] Centre for Computer Studies, Sabaragamuwa University of Sri Lanka, Sri Lanka.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJRCOS/2024/v17i3422

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/112361

Original Research Article

Received: 16/11/2023
Accepted: 22/01/2024
Published: 25/01/2024

ABSTRACT

Identification of the learning style of the students in the teaching and learning environment plays a significant importance in improving both teaching and learning perspectives. The intension of the research was to investigate about applying the Machine Learning Techniques for identification of the Learning style of the students in online learning environment based on the Felder Silverman Learning Style (FSLSM) identification model. The significance of this experiment is that the proposed methodology considers the combination of access frequency (*f*) of course materials and total time (*T*) students spent on each course activity. For data collection process, it was designed reusable plugin for the Moodle for time tracking. Real-time dataset was prepared using three course

**Corresponding author: E-mail: salinda@ccs.sab.ac.lk;*

modules designed according to the features of FSLSM model and the features for analyzing the students according to the FSLSM, it was selected seven criteria, and the features were validated using Pearson Correlation Coefficient method. These course modules were enrolled 150 students per each module. Once the data set was prepared, the data set was preprocessed and applied Five Supervised Classification Machine learning algorithms as Decision Tree, Logistic Regression, Random Forest, Support Vector Machine and K-Nearest Neighbors algorithm. The models were evaluated using Accuracy, Precision, Recall and F1 values. Over the five algorithms the Decision Tree classifier algorithm performed with best average accuracy with 93.5% for Input, 86% for Perception, 89.5 for Processing and 94% for Understanding dimension. The models were validated using the K-fold Cross validation and Standard Deviation values. Mean Squared Error, Bias and Variance values were considered the evaluation of underfitting or overfitting context of the model. To parameter optimization, the Grid Search Methodology was applied to find the best combination of criterion for the model. Finally, an application was developed for Identifying the Learning Style of the Students using the designed Machine learning model. The Consistency of the ML Model based on the Decision Tree classifier algorithm were evaluated using the results generated through developed application and the results suggested that consistency for taught machine learning algorithms is often between 85% to 95%, which is an acceptable range. The results generated by the application for identification of the learning style suggested the combination of learning style for particular students sample as Global-Mild, Visual- Strong, Sensing- Moderate and Reflective-Strong. Identification of these combination of learning style assist for teachers by giving an insight which components of the learning contents should be improved in course designing process.

Keywords: Machine learning; FSLSM; learning style.

1. INTRODUCTION

Integration of E-Learning into education has made many positive changes for improving teaching and learning in the higher education sector [1]. One of the main objectives of systematic education is to develop successful learning mechanisms for the lifelong learning process [2,3] constant learning process is always decided by the way that students acquire the learning contents and how the information is retained with them. Therefore, in the teaching process, one-size fits all method always not success [4] as the learners have their own way of learning styles. Identification of the learning style of the learners plays a pivotal role in the teaching and learning process.

Learning is a necessary human activity that can be shaped and developed both informally and systematically. According to research [5] The teachers have their unique teaching style, and every student has a unique learning style. Students learn and absorb information through hearing, observing, exploring, applying logic, and other methods. Instructors employ a range of instructional techniques as well, such as problem-solving techniques, delivering lectures, and perform with demonstrations. When there is a mismatch between the teaching methods of educators and learners, learners lose interest in the subject and grow bored. Identifying this gap give insight enables them to create courses and content that are targeted to their students' requirements, resulting in a more successful and adaptive teaching experience.

The phrase "learning style" refers to the various techniques that people adopt when participating in educational activities such as reading or studying material offered by instructors or mentors. The public has generally reacted positively to the concept of learning styles [6] occurs because of the learners' unique interactions, abilities, and experiences. Learners can achieve their objectives more precisely with personalized learning models. Identifying each student's learning style is the most important step in personalized learning. To determine the learning style, various learning style models are utilized. Various learning style models, such as Peter Honey and Alan Mumford's model, David Kolb's model, VARK model by Neli Fleming's NASSP model, are used to identify learning styles. Gardner's Theory of Multiple Intelligence, Anthony Gregrorc's Model, and the Felder-Silverman Learning Style Model [7]

This research proposed a methodology for identifying the students' learning style in the online learning environment using Machine Learning (ML) techniques. For this, Moodle was used as the online learning platform for planning the research design and Felder Silverman learning style model (FSLSM) was used as the model for learning style identification and the designing of the learning contents. The research was conducted with three course modules designed in the online learning platform and self-generated dataset was prepared by enrolling the students who are following BSC (Hons) in Information Technology and Information Systems Degree programme at General Sir John Kotelawala Defence University, Sri Lanka.

1.1 Background

Identification of the learning style of the students in the distance learning [8]or online learning environment is one of the focused research areas in the education field [9], These studies focus on identifying the various methodologies for detecting the learning style of the learners and proposing mechanism for improving the teaching and course content design process which tailored with the individual learner's preferences of the learning. Analysis of the student's behavior in the online learning environment and finding mechanisms to measure and determine the actively participation [10]to the course activities and identification of the learning styles of the learners assist in designing the course contents in a more personalized manner. ML as a technology has given tremendous assistance to the analysis of the data related to the leaning preferences and in predicting the Learning Style.

1.2 Research Problem

Since technological advancements have made virtual education more accessible than ever before, researchers are focusing on how to best teach online classes and personalize the educational experience for each individual student. As a result, more research is needed in this area so that teachers can better understand how to approach teaching online courses while yet providing individual attention to each student. The theoretical challenges behind the approaches of analyzing generic Moodle logs and accessing course activity to determine a student's learning style are complicated. Students, for example, may obtain study reference material without necessarily having a predisposition or bias towards that content. This implies that tracking these log hits on a single occurrence may not be a true indication of their actual preferences, and so their learning styles cannot be assessed accurately based on this data alone.

Researchers should look beyond just accessibility features in Moodle logs to gain more insight into student preferences; they should also consider how much time a student spends on certain activities, as well as whether they return to the same activity or content area within the online environment multiple times. This type of study, according to FSLSM, can provide valuable information on what kind of resources students prefer when engaging with online materials, which can then assist shape decisions about teaching tactics customized precisely to those learners' needs and interests.

1.3 Research Significance

As discussed, there are many experiments undergoing to detect the learning style of the students in the online environment using FSLSM model, but the researchers have not considered both the frequency of access as well as the total time spend on each activity as a combine method. The proposed model considers both the frequency with which the course contents are accessed as well as the time spent on each activity and will provide more descriptive insight into learning style identification. And in the methodological approach, a reusable plugin for the Moodle platform was designed which provides the facility of tracking the total time a user spent on each activity accessed. Moreover, reusable FSLSM questionnaire plugin was developed for the collecting of the responses of the Indexed Learning Style Questionnaire. At the final stage, a new application was developed which provides the facility to identify the students' learning style based on the machine learning models developed in the research process.

1.4 Felder Silverman Learning Style model (FSLSM)

Felder and Silverman's Learning Style approach (FSLSM) is a popular approach for understanding students' learning styles [11] .It differs from other models in that it divides learners into four dimensions rather than just a few. This enables for a more in-depth examination of individual student learning preferences and inclinations. Four dimensions of the FSLSM are as follows,

- Active and Reflective style of learning
- Sensing and Intuitive style of Learning
- Visual and Verbal style of learning
- Sequential and Global style of learning

The first dimension looks at how active or reflective the student tends to be when they are engaging with material; this helps teachers understand if they need to provide more guidance or allow them space to explore on their own. The second dimension examines whether the student has a preference towards sensing information through tangible experiences or prefers intuiting abstract concepts; this can help inform instructional strategies such as providing hands-on activities versus theoretical discussions. The third dimension looks at visual versus verbal approaches, while the fourth considers sequential versus global thinking processes - both important considerations regarding how best an individual should be taught certain topics and skillsets.

In a summary, then, Felder & Silverman’s Learning Style Model provides valuable insights into all aspects of an individual’s approach towards education by looking beyond simple groupings based on broad characteristics like age or gender allowing us better tailor teaching methods accordingly so as get the maximum benefit out our instruction efforts.

1.5 Related Works

The study [12] presented information about using Moodle logs and machine learning approaches to identify students' earning styles in an online learning environment. The courses were graded based on three levels of completion: 40%, 60%, and 80%. They have picked forum posts, quiz access, and conversation logs as factors. Machine learning classification algorithms such as Random Forest (RF), Logistic Regression (LR), Decision tree (DTC), (Deep Learning (DL), and Support Vector Machine (SVM), Gradient Boosted Trees (GB), and Naive Bayes (NB) were used to identify the learning style. The Random Forest algorithm provided an optimal accuracy of 81% in the evaluation.

This experiment [13] has been done with participation of 35 students following of computer science module, enrolled the academic course contents using Moodle LMS for their data structure course built using the FSLSM model and Moodle logs. All student activities were tracked by the system, allowing any desired features to be extracted at any point during or after the course. At the end of each lesson, there was also a quiz that tested understanding of what had been taught throughout that session, providing invaluable feedback on how well they had learned from their experience with online teaching methods versus more traditional ones used prior to its introduction into classrooms worldwide. The research employed ROV curve-based approaches to identify the learning mode.

The research [7] includes information regarding the use of machine learning classification approaches for identifying students' learning styles, with the learning styles designed according to the FSLSM and Moodle logs used as attributes. With an accuracy rating of 45.55%, Support Vector Machine (SVM) was shown to be the most accurate in predicting learner's dominant intelligence.

This study [14] proposes a novel method for evaluating learners' automatic and dynamic learning styles based on existing literature. The ILS questionnaire and student behaviour on the LMS are utilized in this method to generate learning style findings that may be used as labels in datasets. Three classifications were tested: decision tree, naive Bayes, and K-nearest neighbour; these models were evaluated using Python with sklearn in two ways: an 80:20 train split test and a K-fold 10 cross-validation test.

Another study was undertaken by Ikawati et al. [15] to identify the learning style of students in the Moodle learning environment utilizing Moodle logs as the attributes. 65 pupils were chosen as the research's data sample. They employed Decision Tree and Ensemble Learning as machine learning algorithms for data classification. The results show that the Decision Tree has an accuracy of 92%, while the Gradient Boosted Tree has an accuracy of 88%.

2. METHODOLOGY

This section of the paper discusses the methodological approach of the paper. As shown in the Fig. 1, As the first step of the research, it was conducted a systematic literature review to find the details about the FSLSM model and the attributes to be used in the learning style identification process in the online learning environment. Then, these selected parameters were validated and applied for the gathering of the data. The next most important step was a planning methodology for analyzing the learner behavior of the online learning environment. In the proposed new methodology, the combination of both access frequency and the total time spent on each activity is considered. For tracking of access frequency, it was used default Moodle logs and for the total time tracking it was designed a new reusable plugin as there is no existing mechanism for counting the total time for each activity in the Moodle.

Then, collected data was preprocessed and trained with supervised classification ML algorithms. After applying the model evaluation methodologies using performance metrics, the best performing model was selected over them and applied the hyper parameter tuning for getting the optimal performance from the best model. Then the students were evaluated by conducting assessments in order to make sure the students engaged with the course activities and their grades were recorded. As the final step separate application was developed for predication of the Learning style of the students in the online learning environment according to the FSLSM learning style model and as another result, it was able define students' performance predication mechanism [16] considering the grades they earned for their assessments which is not described in this research.

2.1 The Dataset

To collect data required for the analyzing the behavior of the learners, it was designed three course modules as IT 1113 Fundamentals of DBMS, IT 2093 Data Structures and Algorithms and IT 3052 Programming Frameworks. Of these three courses, two were used for the compression of the models' accuracy, and one for the measurement of the model's consistency. Each course enrolled 150 students who are following BSc. (Hons) in Information Technology degree programme.

To collect the data the first step was to select the parameters that satisfy the requirements of the FSLSM. As shown in Table 1 seven selected parameters were used for the research process and both frequency and total time were considered. The parameters were validated using Pearson's Correlation technique [17] for feature validation and the results were analyzed using the Heatmaps [18].

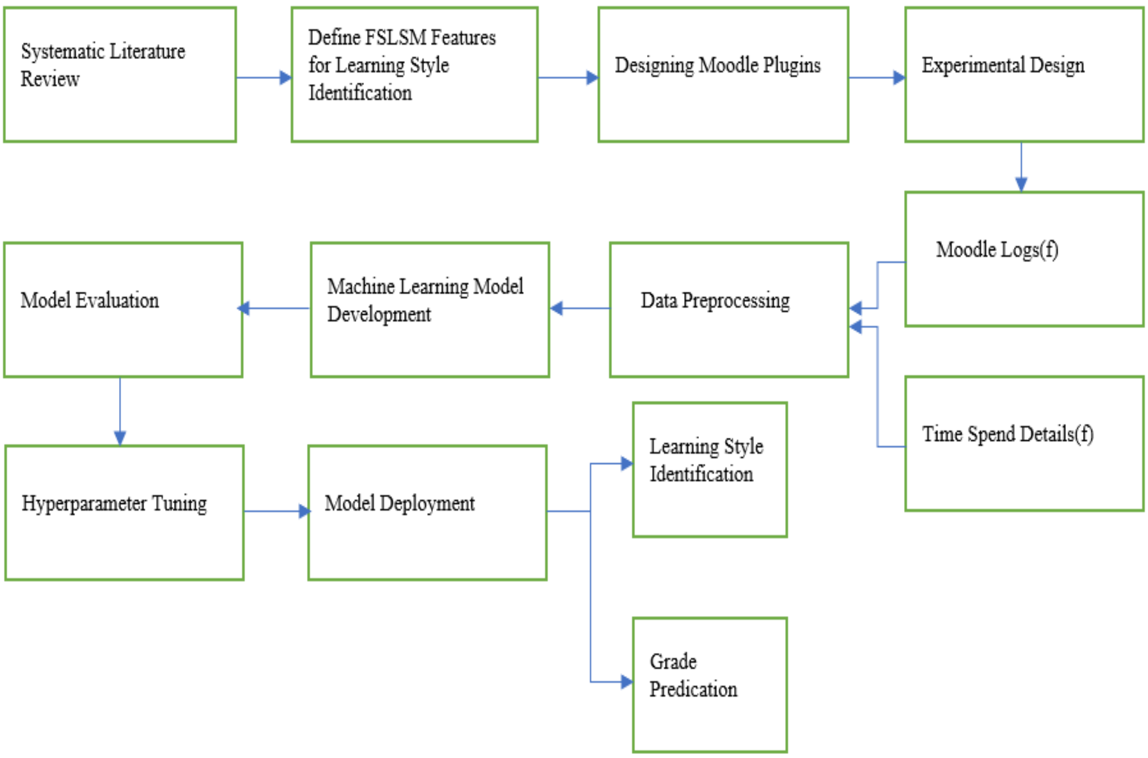


Fig. 1. Main Research steps

Table 1. Feature selection

Dimension	Learning Activity	Learning Style	Expected Behavior	F/T
Processing	Forum posts	Active	Post/Reply	F/T
		Reflective	View (Minimum visits)	F/T
	Exercises	Active	Visit	F/T
		Reflective	No Visit (Minimum visits)	F/T
Input	Videos/Chars/Images/Graphs	Visual	View	F/T
	(V_C_I_G)	Verbal	No View	F/T
	Notes	Visual	View	F/T
		Verbal	No View	F/T
Perception	Contents	Sensing	Concrete Contents	F/T
		Intuitive	Abstract Contents	F/T
Understand	Course Outlines	Global	Course Outlines View	F/T
		Sequential	No View Visit (Minimum visits)	F/T

2.2 Performance Metrics

2.2.1 K-Fold Cross validation

K-fold cross-validation is a validation technique used for further evaluation of the model's accuracy. In here, K-fold cross-validation gives an insight of how the model will behave for a scenario of new unseen data [19].The Standard Deviation (SD) value generated with the cross-validation process gives an insight that how the values of the dataset clustered closely around the mean value [20].The lower SD values implies that the model datapoints have a good similar relationship with each other.

2.2.2 Mean Squared Error (MSE)

Mean Squared Error (MSE) can be called a methodology used for the performance evaluation of an ML model. It generates the performance results by measuring the average squared difference value of the target variables. A lower MSE value indicates the model is better in the performance implying model predications are closer to the actual predication accuracy values [21].

2.2.3 Bias value

The term "bias" in the context of MSE refers to a model's systematic mistake or tendency to consistently underpredict or overpredict true values. It calculates how far the forecasts are, on average, from the actual numbers. A model with a high bias oversimplifies the underlying relationships in the data, which can lead to underfitting. It consistently creates errors and fails to capture the intricacy of the data [22].

2.2.4 Variance

The unpredictability or instability of a model's predictions across multiple training sets is referred to as "variance" in the context of MSE. A model with a large variance is highly sensitive to training data and may overfit. It picks up on noise and unpredictable fluctuations in the training data, resulting in poor generalization to new data. MSE's variance component represents the average squared difference between anticipated and actual values. A lower variance means that the model's predictions are more consistent across training sets [23].

2.3 Time Tracking Plugin for Moodle

This is one of the major contributions of the research in deciding the learning style of the students. In Moodle platform, there are mechanisms for counting the time spent completing an assignment like tasks, but there is no methodology for tracking how much time a user spends on all the activities relevant to the course module. For this, it was designed a distinct new plugin for counting time. A snapshot of the UI is displayed.

The plugin runs on every course activity page and uses JavaScript libraries to track how much time a user spends on each course activity that has been linked with the plugin. When a query is made, the recorded time is saved in the Moodle database and may be accessed using PHP. The Moodle format assigns each user a unique ID, making it easy to track how much time each student spends on an activity. The time is captured in milliseconds and then converted to seconds. Teachers and course administrators with privileged access can view and download time-spent statistics in CSV format. Students have no access to the collected data.

2.4 Indexed Learning Style (ILS) Questionnaire plugin

As shown in the Fig. 3, a plugin for ILS plugin was developed for the Moodle which is reusable. The students can access this and provide their responses for the ILS questionnaire. ILS questionnaire is an already accepted questionnaire for identifying the learning style of the students according to the FSLSM [24]. This consists of forty-four questions that represent each dimension by eleven questions. In this plugin, the conventional ILS questionnaire has been converted to a computerized format, and the purpose is that these responses were used in the labeling process in preparing the dataset.

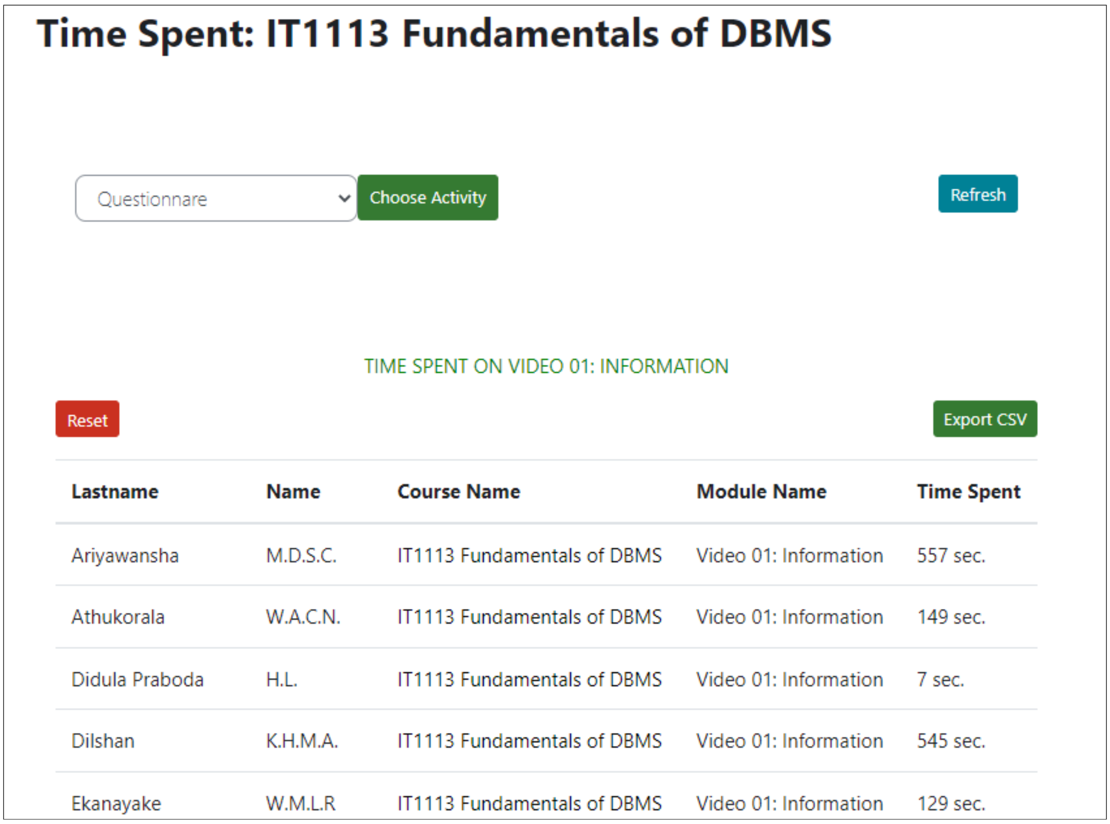


Fig. 2. Time tracking plugin

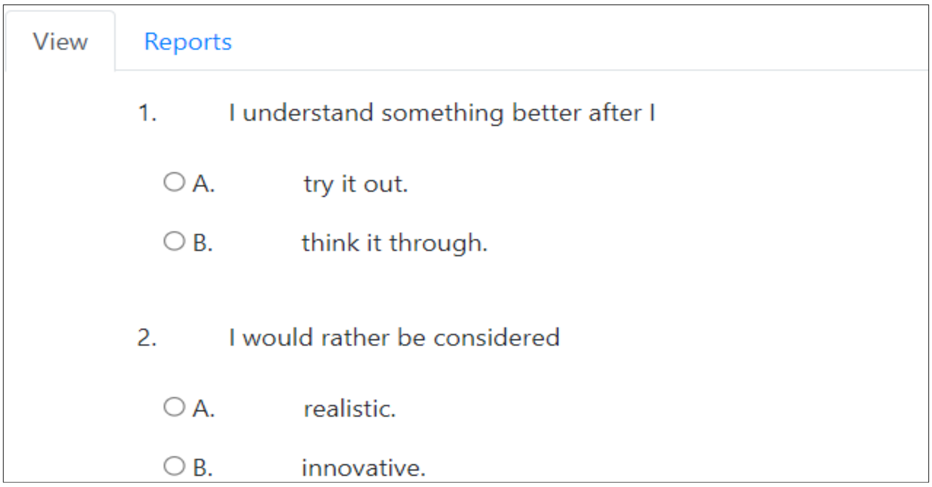


Fig. 3. ILS Plugin

2.5 Moodle Course Designing

For analyzing the behavior of the learners in online learning environments according to the parameters of FSLSM, three courses were designed. In the course designing process, each course was included course materials are shown in the diagram in the Fig. 4, It was used forum posts, exercises, videos, charts, images, graphs, notes, theoretical concepts, conceptual maps according to the rule classifications of the FSLSM.

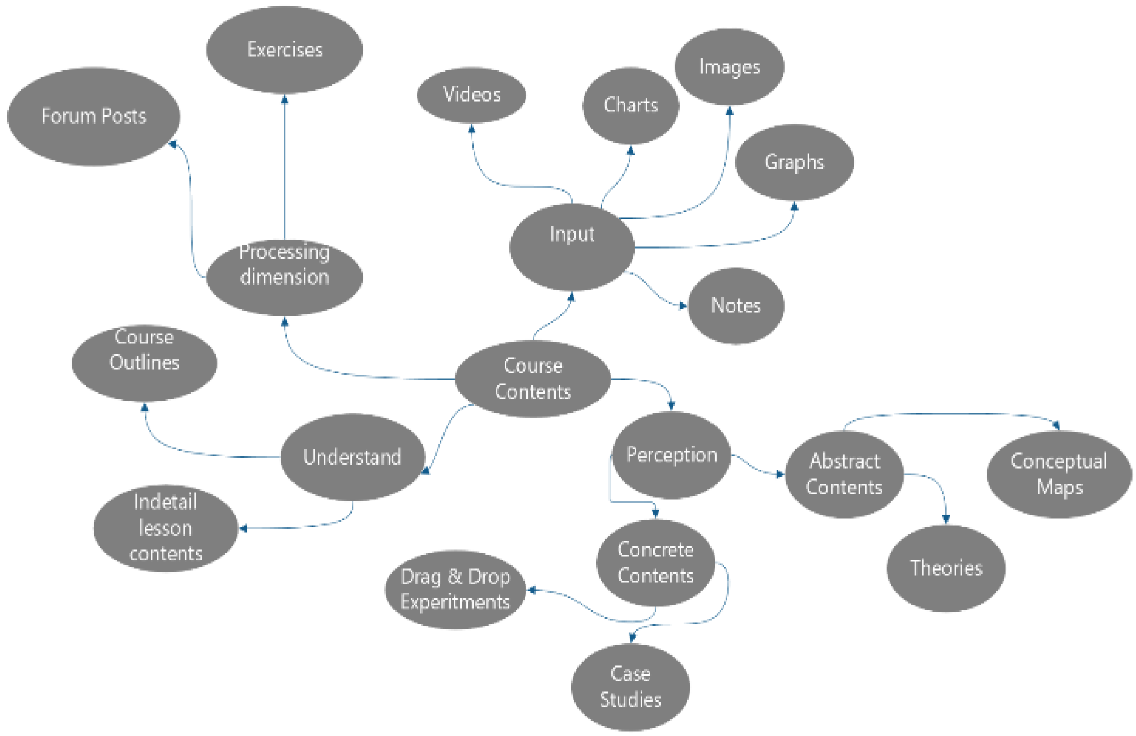


Fig. 4. Course Design Map

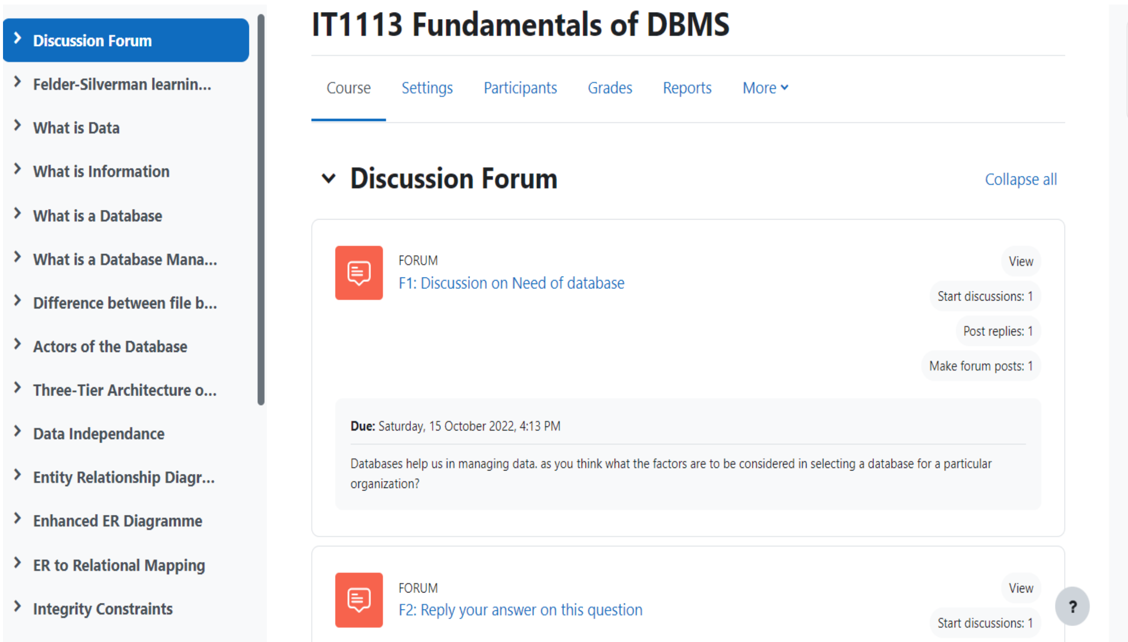


Fig. 5. Course Designing in Moodle

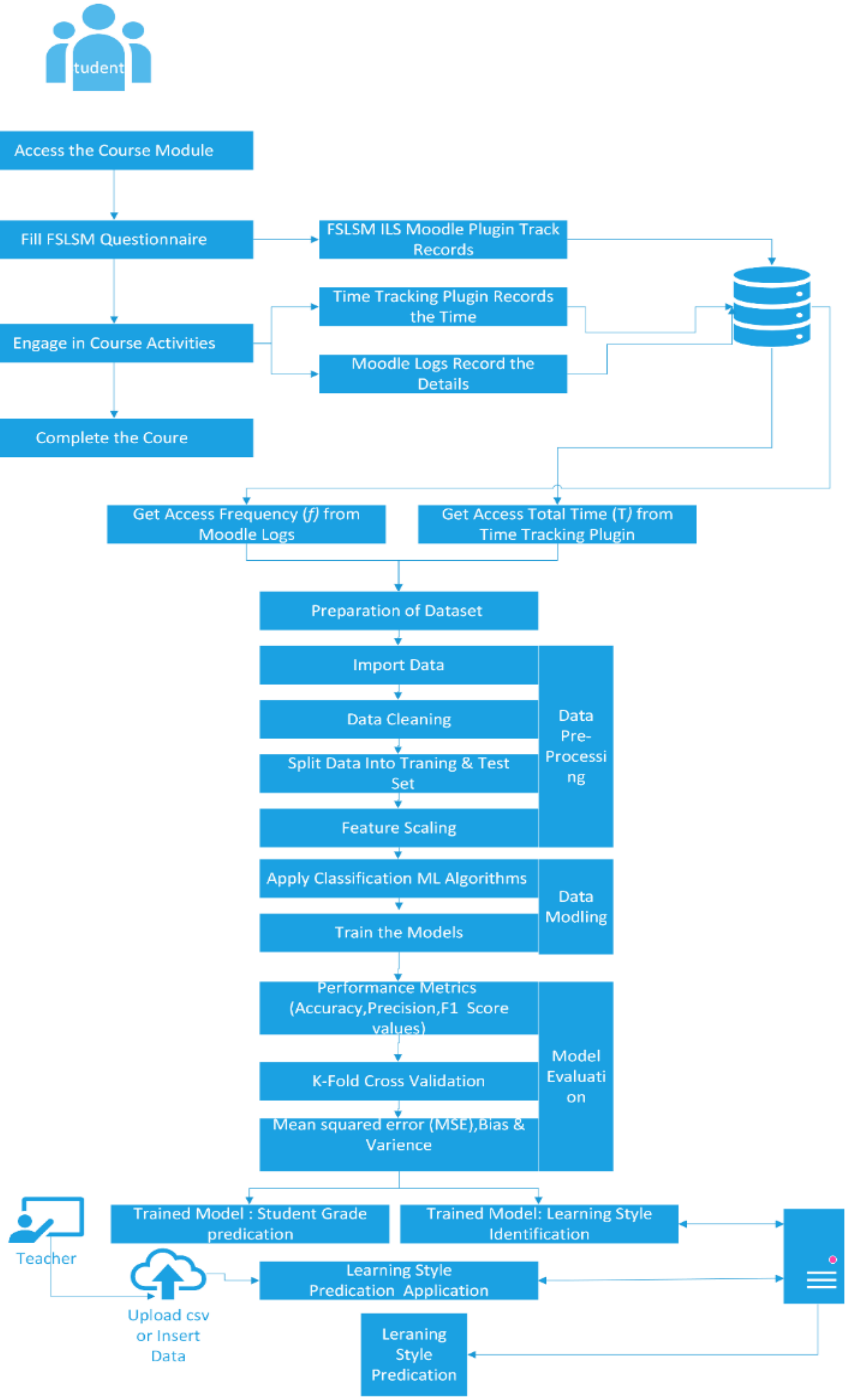


Fig. 6. Experimental design

Fig. 5 shows sample interface of one of the courses designed for the experiment process. As the first steps it was decided the major topics to be included in the module as per the course guidelines. And the next important part was to design the contents tailored with the parameters of FSLSM of analyzing the learner behavior in the online learning environment. For each topic, learning activities were designed representing each major four dimension of the FSLSM model as Processing, Input, Perception and Understanding. If consider a one topic related to the module, it was contained the course materials to represent the contents for forums, videos, images, conceptual maps, abstract contents, concreate contents and selected other features.

2.6 Experimental Design

As shown in the Fig. 6, Once the Students access the course module. First, the students are informed to fill in the ILS questionnaire. These responses are collected for the preparation of the dataset for the ML models. Then students can be engaged with the course activities and will access the course materials as per their preferences and study the modules. Finally at the end the students are evaluated by conducted assessments and recorded their performances. Once students complete the course module, their behaviour in the learning environment is tracked by Moodle logs already available as a default facility of the Moodle and the plugin designed for tracking the total time. In this the frequency of access to the course activities and total time spent on each activity is collected.

Once the data is prepared, the data is going under preprocessing stages for data cleaning, splitting data to training and test. And feature scaling. Then Data modeling is done by applying the Machine learning algorithms. For this it was used five Supervised Machine learning Algorithms as Decision Tree Classifier, Logistic Regression Classifier, Random Forest Classifier, Support Vector Machine Classifier and K-Nearest Neighbors Classifier. Then the models were validated using Accuracy, precision, Recall, F1 Values as well as K-fold grid validation was done to validate the model by analyzing how the model will behave for a new scenario. In here Standard Deviation also was considered. Then the model was tested with Mean Squad Error (MSE) Bias and Variance values also were calculated to measure whether the model is having under fitting or over fitting situations. Above mentioned methods were applied for models tested for the two course modules as Fundamentals of DBMS and Data Structures and Algorithms and compared the average accuracy of the models.

After selecting the best accurate model among the tested algorithm, Grid Search Methodology was applied to find the best combination of criteria for optimizing the model. Once the model is optimized, an application was developed for identifying the students' learning style based on the developed model. An interface of the developed application is shown.

AS shown in the Fig. 7, once set of data uploaded to the developed application, it will predicate the learning style of the students. Then the teacher can provide feedback to the students according to the learning style provided for the students.

The interface is shown in Fig. 8.

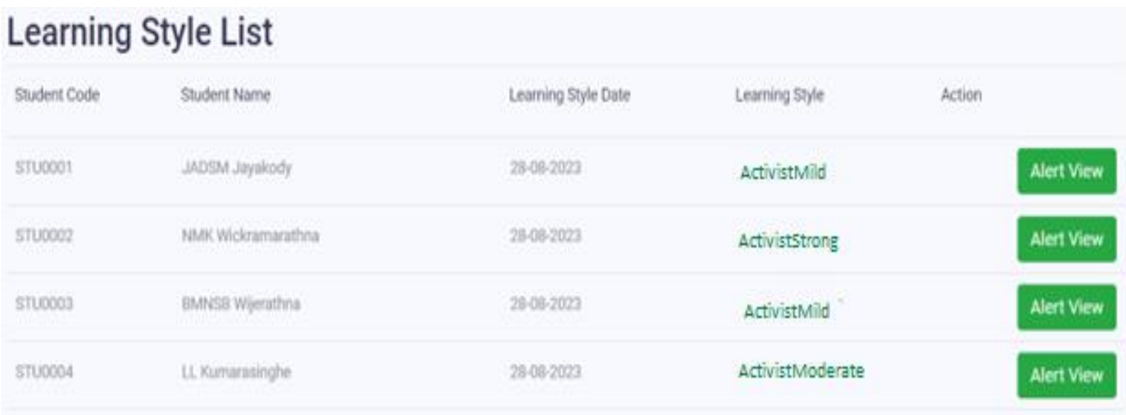


Fig. 7. Learning style identification

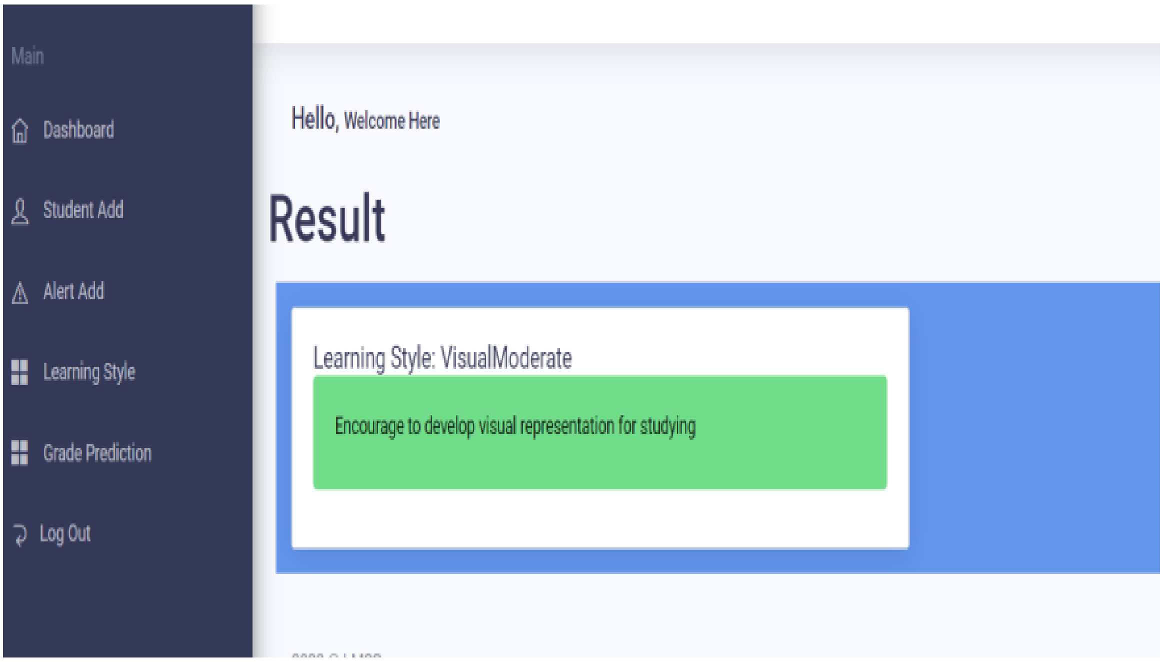


Fig. 8. Providing Recommendations

Once the application was developed, the consistency of the models was tested using the other course module Programming Frameworks. In this it was compared the learning style generated by the developed application based on the ML model and the learning Style identified through the ILS questionnaire.

3. RESULTS AND DISCUSSION

This section of the paper discussed the results generated in the feature validation process and the ML model development process.

3.1 Feature Validation

The feature validation process was conducted to identify whether there is a relationship between the selected features for analyzing the behavior of the learners in the online learning environment. It was selected seven features for the experiment and for each feature, it was considered both frequency (f) and the time (T) spent on each activity. The purpose of the feature validation in before continuing the experimental process, it is required to confirm the existence of the relationship between selected attributes and the dimension, and this can be used to identify any other relationships among attributes as well. For this it was used Pearson Correlation Coefficient Value, and the results were visualized using Heatmaps.

3.1.1 Input dimension

The selected features for Input dimension are as Videos V_C_I_G (f), Notes (f), Videos V_C_I_G (T) and Notes (T). Here the feature V_C_I_G (f) has correlation value of 0.56 which is considered as a good relationship between the dimension and the feature. The access frequency of Notes (f) has correlation of 0.6 which is a good relationship. And time spent on V_C_I_G (T) has good correlation value of 0.51 and. Time spent on Notes (T) is 0.89 which is considered as a stronger relationship. This implies that selected features for the Input dimension have a good relationship with the dimension.

3.1.2 Processing dimension

The selected features for the Processinf dimension are as Forum (f), Exercises (f), Forum (T) and Exercises(T). for the frequcent of access Forum(f) gives correlation coefficient of 0.82 which is consdierd as a stronger relationship. The feature Exercisse (f) also have higher coeficient value of 0.87 with a stronger relationship.Time spent on activities the feature Forum (T) has a value of 0.69 which implies there is good relationship and Exercises(T) show 0.82 of stronger value.As a summery, selected attributes have a good relationship with the dimension.

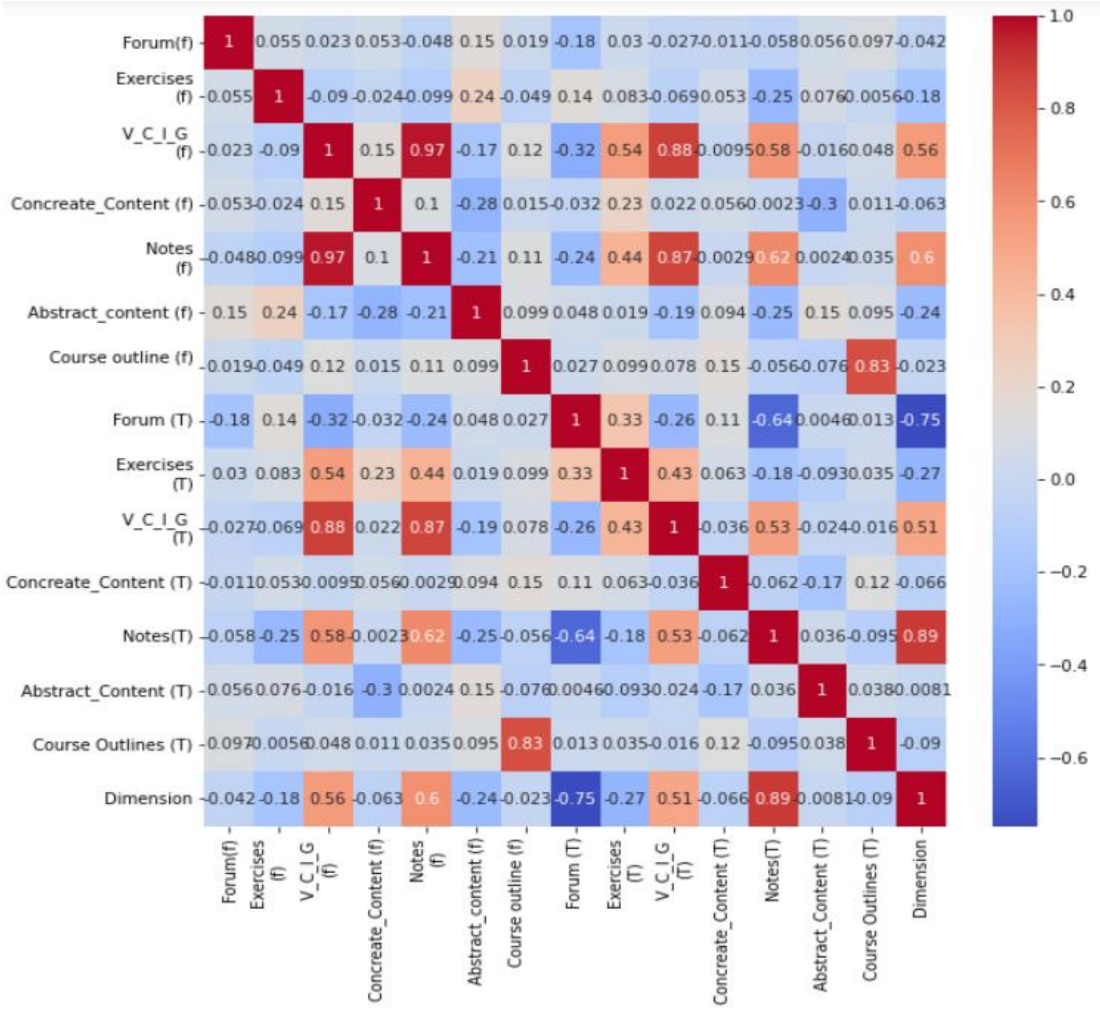


Fig. 9. Input dimension features

3.1.3 Perception dimension

The selected features for the Processing dimension as Concreate (f), Abstract (f), Concreate (T) and Abstract (T). The frequency of access Concreate (f) shows coefficient value of 0.62 which is a good relationship and Abstract (f) gives a value of 0.61 which is a good value implies there is a good relationship. The time Spent Concrete (T) gives value of 0.61 which is good value and Abstract (T) gives a value of 0.54 which implies there is good relationship.

3.1.4 Understanding dimension

As shown in Figure 12, the selected features for the understanding dimension are as Course Outlines (f) and course outlines (T). The feature Course Outlines (f) frequency of accessing the Course outlines has a correlation coefficient of 0.88, indicating a strong positive association between the dependent variable and the feature variable, as seen in the Heatmap. The correlation coefficient value for the feature variable Course Outlines (T), total Time is 0.88, indicating a strong positive association between the Dependent variable and the feature variable.

Finally, for the experiment, these attributes can be used to analyses the behavior.

3.1.5 Machine learning model evaluation

This section of the document discusses the ML model evaluation. Here for each four dimensions of FSLSM, the evaluations were done to find the most appropriate model. For each perspective five ML supervised classification algorithms were tested. In here, the experiment was done for two course modules. The evaluation results are discussed here under each main four dimensions.

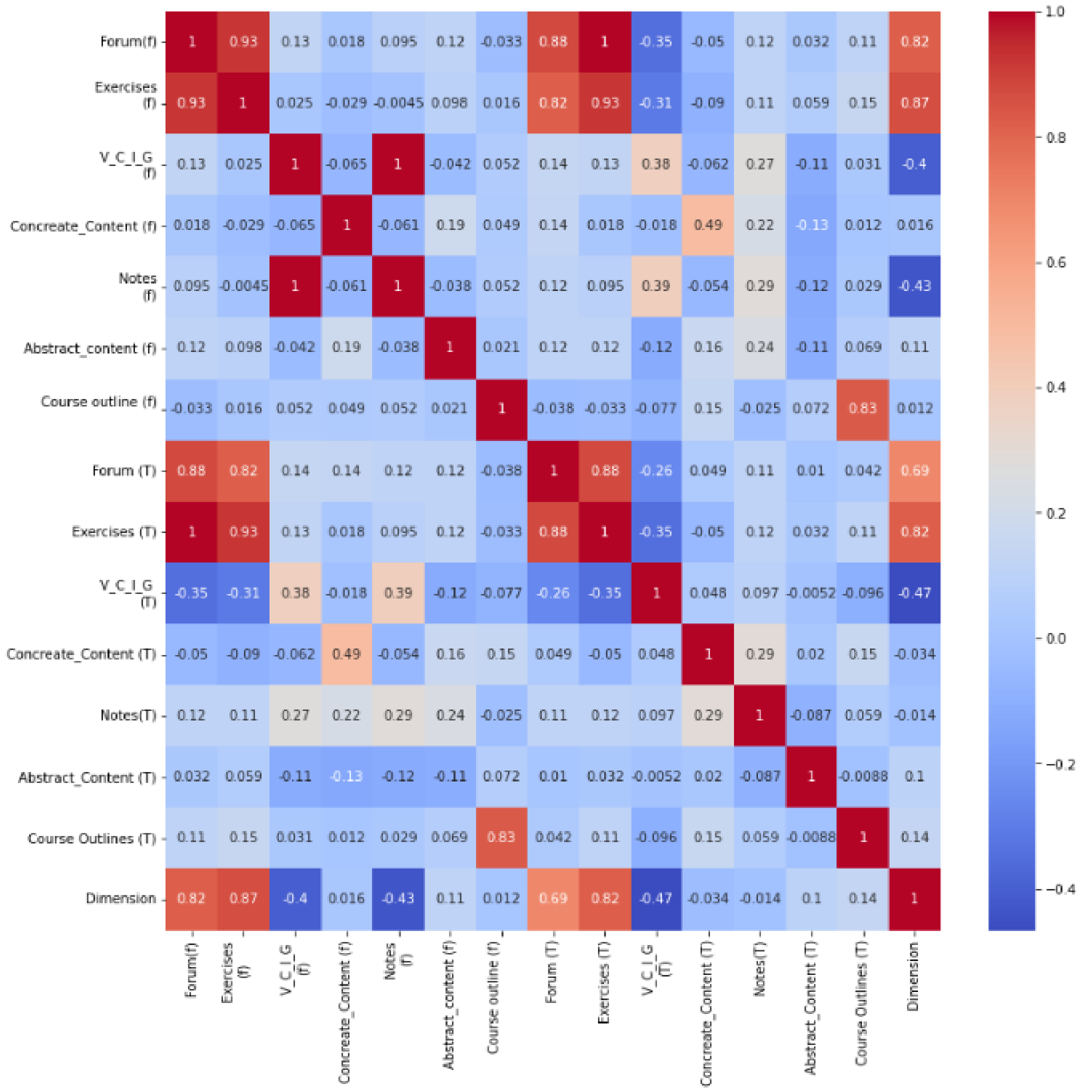


Fig. 10. Processing dimension features

Table 2. Results of Fundamentals of DBMS Module - Input Dimension

Dimension Algorithm	Input Dimension			
	Accuracy	Precision	Recall	F1
Decision Tree	95%	82%	80%	77%
Logistic Regression	61%	34%	42%	37%
Random Forest	85%	67%	73%	69%
Support Vector Machine	71%	56%	54%	55%
K-Nearest Neighbors	60%	36%	43%	38%

3.1.6 Input dimension

Here it shows the results generated for Fundamentals of DBMS module. The Decision Tree algorithm was performed with the highest accuracy of 95%. The precision is also relatively good at 82%, implying that a considerable proportion of the predicted positive cases were correct. The recall is good at 80%, indicating that the model can identify most true positive cases. Finally, the F1 score is 77%, indicating a good combination of precision and recall. Overall, the model appears to perform well on the given task, with a reasonable balance of precision and recall. The results generated for K-fold cross validation is 96% with a SD value of 10. This implies that the model is at a good level of make predication for new scenario and most of the data points are around the mean value. The MSE value is 0.011, bias is 0.008, and variance is 0.004. The model has relatively low values of MSE implies that model is better fit and moderate variance implies model is not overfitting or underfitting.

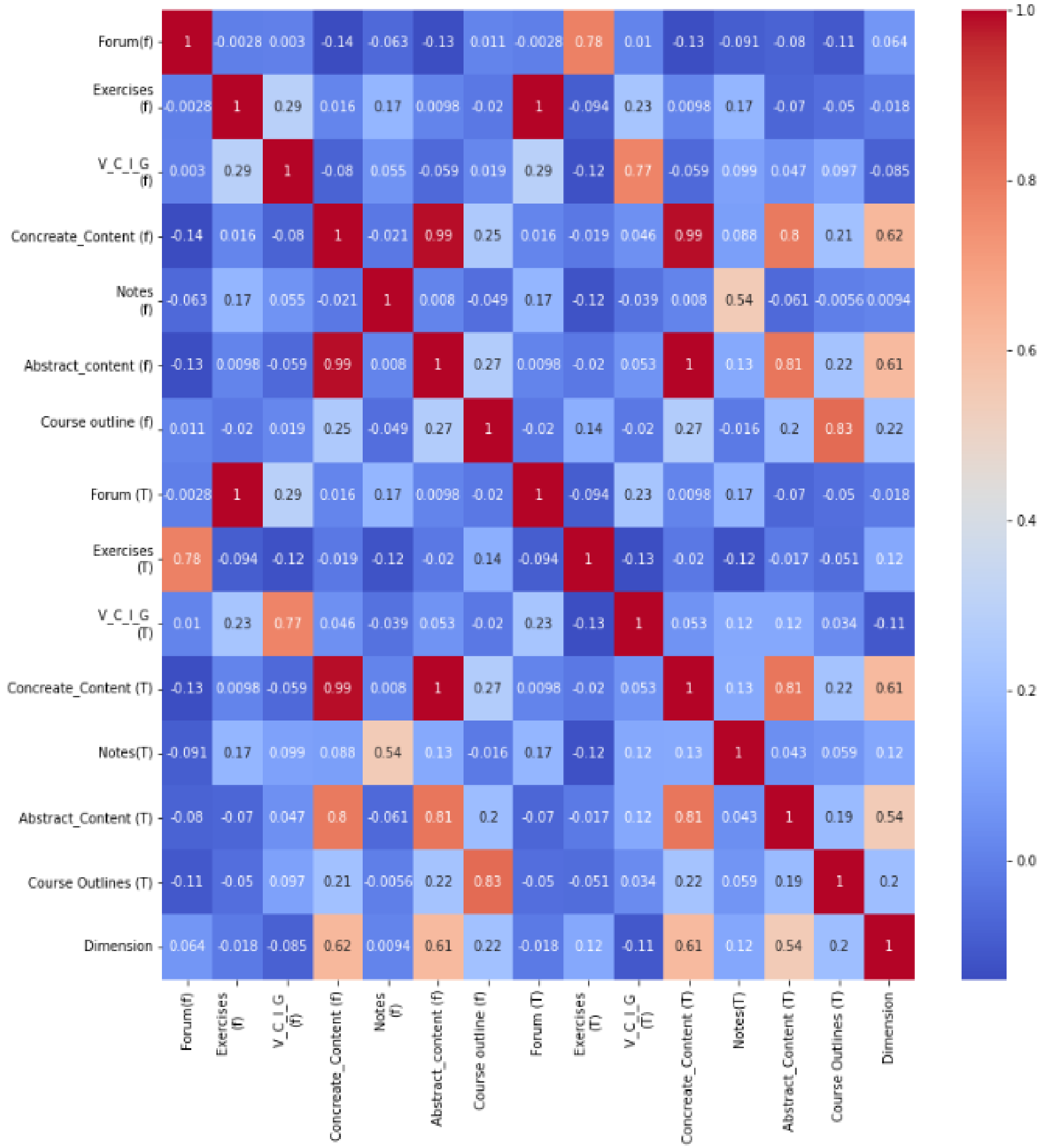


Fig. 11. Perception dimension features

Table 3. Results of Data Structures Algorithms Module - Input Dimension

Dimension Algorithm	Input Dimension			
	Accuracy	Precision	Recall	F1
Decision Tree	92%	81%	78%	75%
Logistic Regression	58%	32%	41%	36%
Random Forest	86%	69%	74%	65%
Support Vector Machine	69%	54%	53%	50%
K-Nearest Neighbors	64%	39%	44%	36%

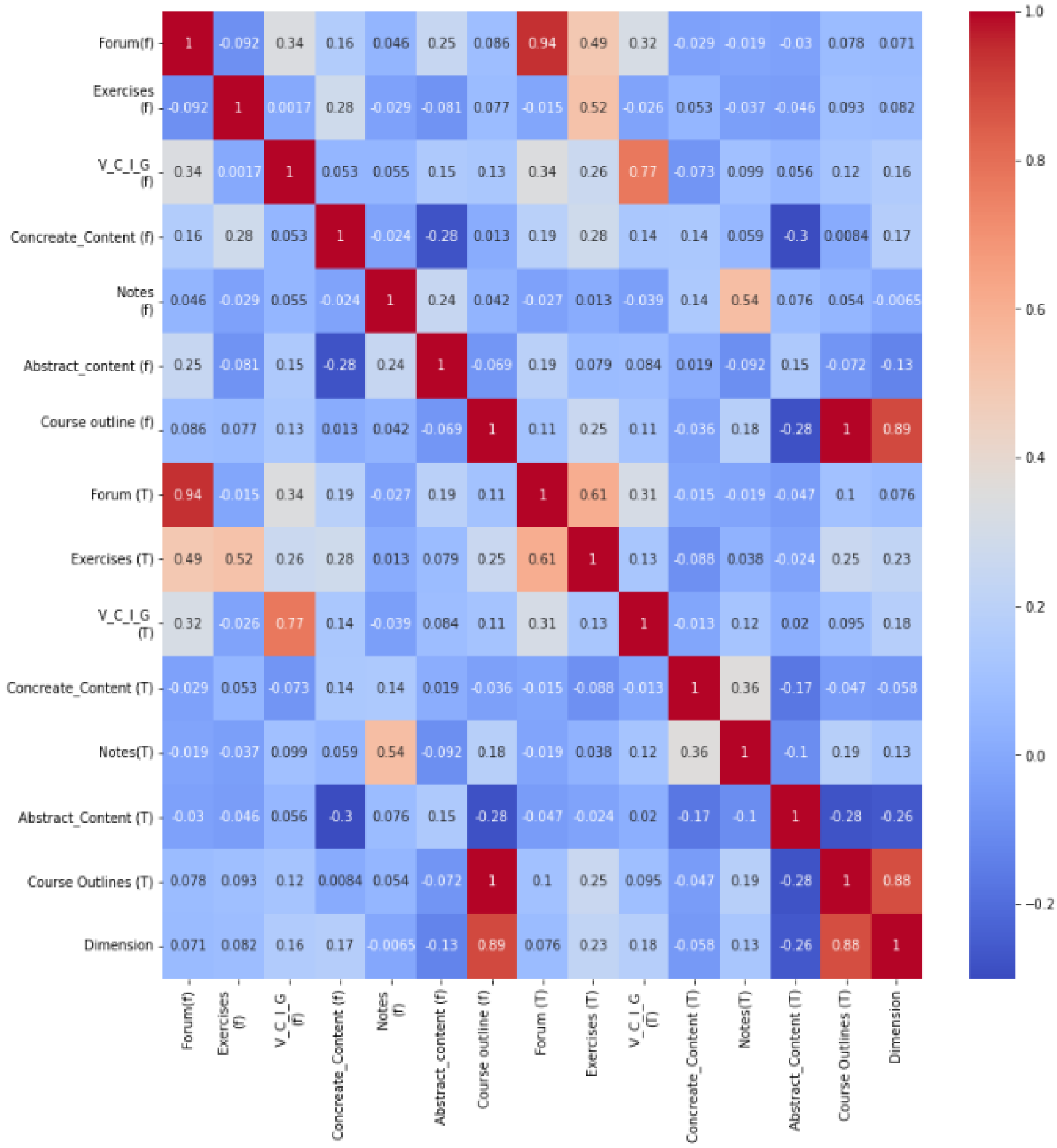


Fig. 12. Understanding dimension features

For this, The Decision Tree algorithm has performed with 92% accuracy. The Precision recall and F1 values are also in a higher percentage which implies that model is in a good level of predication positive true values. Goof F1 value indicates a good combination of precision and recall values. The MSE value is 0.021, Bias is 0.014 and the Variance is 0.006. The lower values of Bias and Variance implies there is no underfitting and overfitting context regarding the selected algorithm model.

Table 4. Results of Fundamentals of DBMS Module -Perception Dimension

Dimension Algorithm	Perception			
	Accuracy	**Precision**	**Recall**	**F1**
Decision Tree	80%	75%	72%	70%
Logistic Regression	44%	49%	36%	35%
Random Forest	61%	50%	41%	44%
Support Vector Machine	57%	51%	44%	44%
K-Nearest Neighbors	66%	47%	51%	49%

3.1.7 Perception dimension

Here, The Decision Tree algorithm has performed with 80% of good Accuracy value. The precision, recall and F1 values are also at a good level. This implies that the Decision Tree algorithm is at a good level of predicating the results. Cross Validation value is 78% which is a good percentage and SD value is 9.72 which is in a good range implies most of data points are around the mean value. The MSE is 0.023, Bias 0.007 and Variance is 0.005. These values are relatively low values that implies model is not having underfitting or overfitting contexts.

The Decision Tree Algorithm has performed with an 85% accuracy which is a good value and Precision, Recall and F1 values are in 78%, 73% and 70%. So, the model is in a good state to make the predictions. And there is a good balance between Recall and Precision values. The MSE value generated for the Decision Tree Classifier is 0.024, Bias is 0.006 and Variance is 0.007. The lower values of these parameters implies that the Decision Tree algorithm-based model is in a good state with no underfitting or overfitting contexts.

3.1.8 Processing dimension

As shown in Table 6, the decision Tree algorithm has performed with 90% accuracy and precision value of 94% recall value of 92% and F1 value of 92%. This result shows the model is in a good state of making predictions and good balance of the recall and F1 values. Cross validation is also 81% which is a higher accuracy that implies model is in good level of make prediction for a new dataset as well. The SD is 8.16 that gives an insight into how most of the data points are around the mean value. The MSE is 0.034, Bias is 0.018, and Variance is 0.007. Hence the model is not giving any insight of underfitting or overfitting.

In the Data Structures and algorithm module The Decision Tree algorithm performed with an 89% which is a higher accuracy. The Precision, Recall and F1 values are also at a good level that implies the model is at a good level for making predictions. The MSE value generated for the model is 0.021, Bias is 0.019, and the variance 0.006. The lower values give an insight that the model is not having any underfitting or overfitting contexts

Table 5. Results of Data Structures and Algorithm Module - Perception Dimension

Dimension Algorithm	Perception			
	Accuracy	Precision	Recall	F1
Decision Tree	85%	78%	73%	70%
Logistic Regression	46%	44%	37%	35%
Random Forest	69%	52%	44%	42%
Support Vector Machine	55%	52%	46%	44%
K-Nearest Neighbors	69%	57%	54%	47%

Table 6. Results of Fundamentals of DBMS Module -Processing Dimension

Dimension Algorithm	Processing			
	Accuracy	Precision	Recall	F1
Decision Tree	90%	94%	92%	92%
Logistic Regression	61%	67%	78%	63%
Random Forest	66%	57%	58%	57%
Support Vector Machine	71%	84%	83%	83%
K-Nearest Neighbors	42%	41%	45%	42%

Table 7. Results of Data structures algorithms Module -Processing Dimension

Dimension Algorithm	Processing			
	Accuracy	Precision	Recall	F1
Decision Tree	89%	84%	80%	79%
Logistic Regression	64%	62%	61%	59%
Random Forest	68%	58%	56%	50%
Support Vector Machine	74%	72%	80%	82%
K-Nearest Neighbors	44%	43%	42%	41%

3.1.9 Understanding dimension

As shown in Table 8, The Decision Tree algorithm performed with 95% higher accuracy. Which is the best level. The Precision, Recall and F1 values are also at a higher level that implies the model is at a good level of making positive predictions. The value given in the K-fold grid validation is 81% which is a good accuracy level which implies the model is at a good level of making predications for a new dataset. The SD value is also 6.14 which gives an insight the data points are around the mean value. The MSE value is 0.056, Bias is 0.042 and Variance is 0.008. The values are in a lower level which gives an insight that there is no underfitting or overfitting context in the model.

As shown in Table 9, the Decision Tree algorithm has performed with the highest accuracy of 93%. This gives an insight into why the model is in a good state of making predictions. The precision recall and F1 values are also in a good state that implies that there is a good balance of Precision and recall values. The MSE value generated for the Decision Tree Algorithm based model is 0.047, Bias is 0.042 and Variance is 0.007. These lower values of Bias and Variance implies there is not any underfitting or overfitting contexts.

3.2 Comparison of Results of Models

The previous section of the paper discussed the evaluation of the ML models for selected algorithms. The purpose of this was to check the behavior of the models for two distinct courses for each of the four dimensions.

Table 9 gives the results generated for Data Structures and Algorithms Module.

3.2.1 Input dimension

As shown in Fig. 13, for both course modules, The Decision Tree Classifier has performed with the highest average accuracy of 93.5% over the other selected algorithms. As the second only Random Forest algorithm is in an average accuracy of 80%. This implies that, For the Input Dimension, the Decision Tree Classifier is performing with the best accuracy.

3.2.2 Perception dimension

As shown in Fig. 14, The Decision Tree algorithm has performed with the best accuracy of average 86%. This implies that over the other compared classification algorithms, Decision Tree classifier is performing with best accuracy.

3.2.3 Processing dimension

As shown in the Fig. 15, over the compared algorithms the Decision Tree classifier has performed best with average accuracy of 89.5%. This implies that Decision Tree Classifier is best performing for the Processing Dimension.

3.2.4 Understanding dimension

As shown in Fig. 16, For the Understanding dimension as well, The Decision Tree Classifier has performed with the best average accuracy of 94%. This gives an insight into whether the Decision Tree Classifier is performing best for the Understanding dimension as well.

Table 8. Results of Fundamentals of DBMS Module -Understanding Dimension

Dimension Algorithm	Understanding			
	Accuracy	Precision	Recall	F1
Decision Tree	95%	83%	80%	78%
Logistic Regression	74%	40%	41%	40%
Random Forest	85%	42%	50%	46%
Support Vector Machine	90%	70%	62%	64%
K-Nearest Neighbors	47%	23%	27%	24%

Table 9. Results of Data Structures and Algorithm Module -Understanding Dimension

Dimension Algorithm	Understanding			
	Accuracy	Precision	Recall	F1
Decision Tree	93%	84%	81%	76%
Logistic Regression	78%	43%	40%	42%
Random Forest	83%	41%	52%	44%
Support Vector Machine	92%	75%	64%	66%
K-Nearest Neighbors	48%	24%	20%	22%

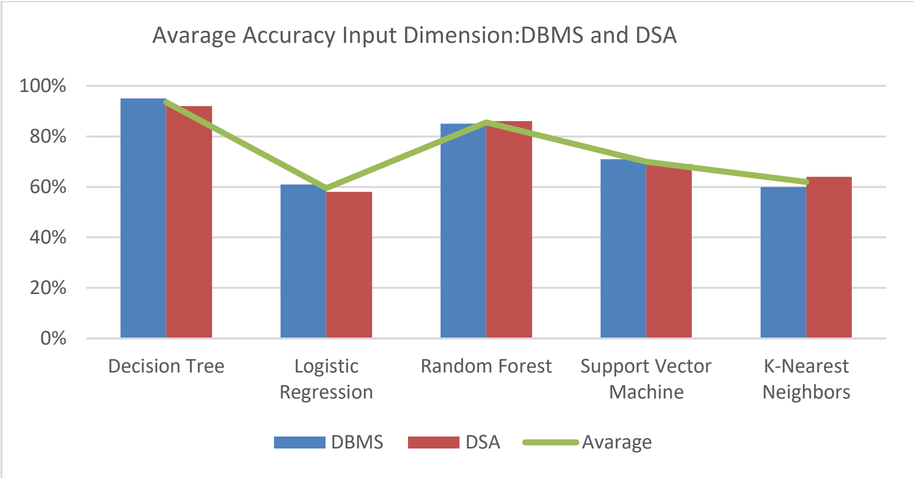


Fig. 13. Average Accuracy: Input Dimension

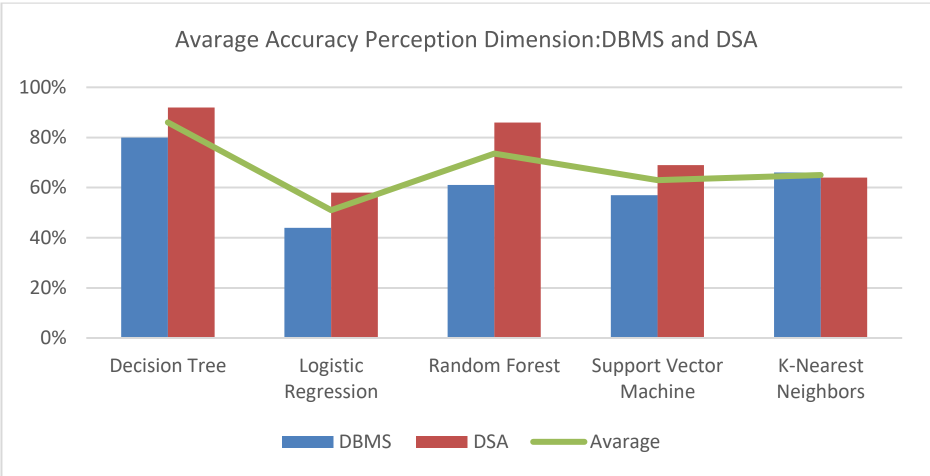


Fig. 14. Average Accuracy: Perception dimension

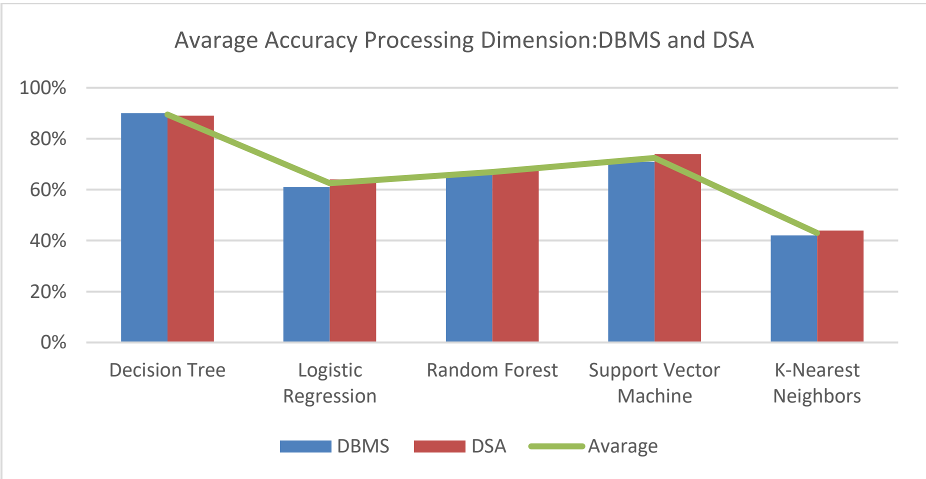


Fig. 15. Average Accuracy: Processing Dimension

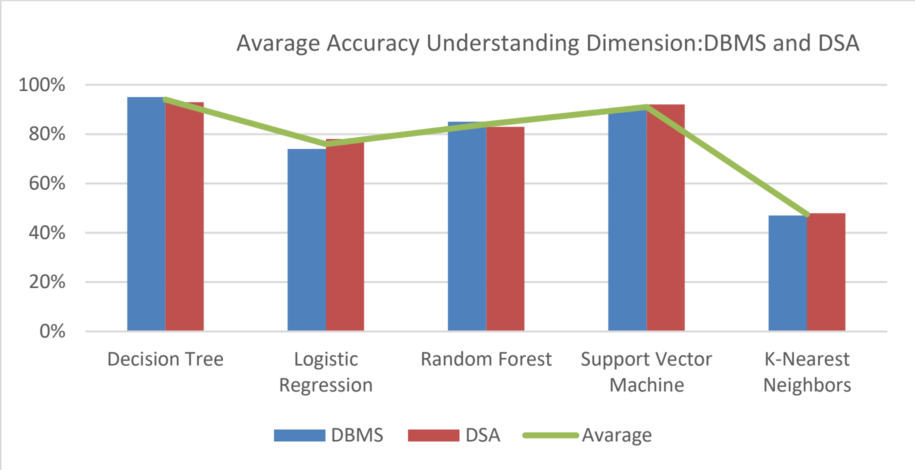


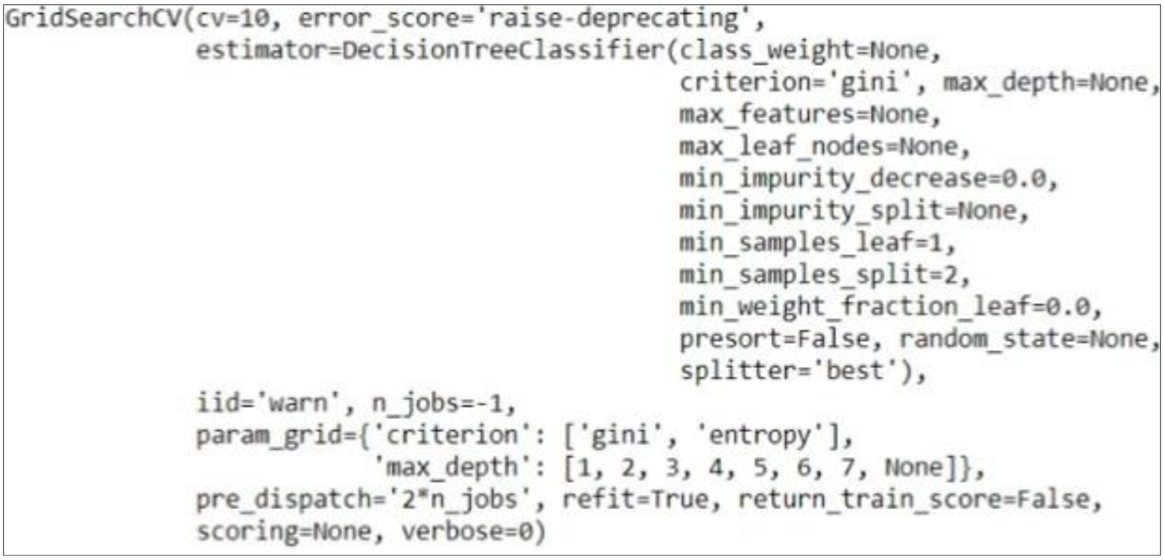
Fig. 16. Average Accuracy: Understanding Dimension

As a summary, for each four dimensions, The Decision Tree algorithm has performed with the best accuracy. Furthermore, when analyzing values generated for the MSE, Bias and Variance, The Decision Tree classifier model has given the lower values by giving an insight that there is no underfitting or overfitting context. Hence for the process of Identification of the Learning Style, The Decision Tree Classifier based ML model was used for each four dimensions.

3.3 Hyperparameter Tuning with Grid-Search

The Grid-Search Methodology finds the combination of best performing parameters for the selected ML model. In here as the research used classification algorithms, it was referred to the scikit-learn library and criteria were selected by applying the Grid- Search Methodology as shown in Fig. 17.

According to the results, the parameter combination for the parameter tuning. {'criterion':gini,'max depth':2}. The gini parameter works well for the gini parameter over the 'entropy' with the max depth of the tree =2. Then these combinations were applied to the ML model designed for the Leaning Style identification process for each of the four dimensions.



```
GridSearchCV(cv=10, error_score='raise-deprecating',
             estimator=DecisionTreeClassifier(class_weight=None,
                                              criterion='gini', max_depth=None,
                                              max_features=None,
                                              max_leaf_nodes=None,
                                              min_impurity_decrease=0.0,
                                              min_impurity_split=None,
                                              min_samples_leaf=1,
                                              min_samples_split=2,
                                              min_weight_fraction_leaf=0.0,
                                              presort=False, random_state=None,
                                              splitter='best'),
             iid='warn', n_jobs=-1,
             param_grid={'criterion': ['gini', 'entropy'],
                         'max_depth': [1, 2, 3, 4, 5, 6, 7, None]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=False,
             scoring=None, verbose=0)
```

Fig. 17. Hyperparameter tuning

3.4 Evaluation of the Consistency of the ML Model

The consistency of the course modules was measured by the results based on the course module Programming Frameworks. The mechanism applied in here was, for the entire 150 students, first, Learning Styles were identified according to the Conventional ILS Questionnaire, and then these generated results were compared with the same module with the results generated through the developed ML model. According to the results shown in Fig. 18, it shows the consistency confidence for each four dimensions as Input, Perception, Processing and Understanding.

When analyzing the behaviour of each line associated to the dimension, it is clear that there are no significant differences between the findings obtained using the ILS questionnaire and the machine learning model. As an average, consistency ranges from 85% to 95%, which is a good level of consistency for taught Machine learning algorithms. As a result, all Machine Learning models built for each dimension are satisfactory.

3.5 Analysis of the Pattern According to the Results

The following digramme shows the summary of the patterns identified for the results.

(Fig. 19), According to the results displayed in the above chart (Fig. 19), there is a rather high preference for the visual Strong sub dimension when it comes to the students' input dimension, which relates to how people prefer to receive information. The perception dimension provides insight into the way individuals choose to receive information through their senses. This pattern indicates that learners strongly like the Sensing mild dimension. According to the findings, there is a comparatively strong preference for the Active Mild and Reflective Strong in the Processing Dimension, which measures how people like to process information. The understanding component of the Felder-Silverman learning style identification model (FSLSM) describes an individual's preferred method of acquiring new knowledge or ideas. According to the results, there is a comparatively strong preference for the Global-Mild dimension. As summary the results shows that the pattern

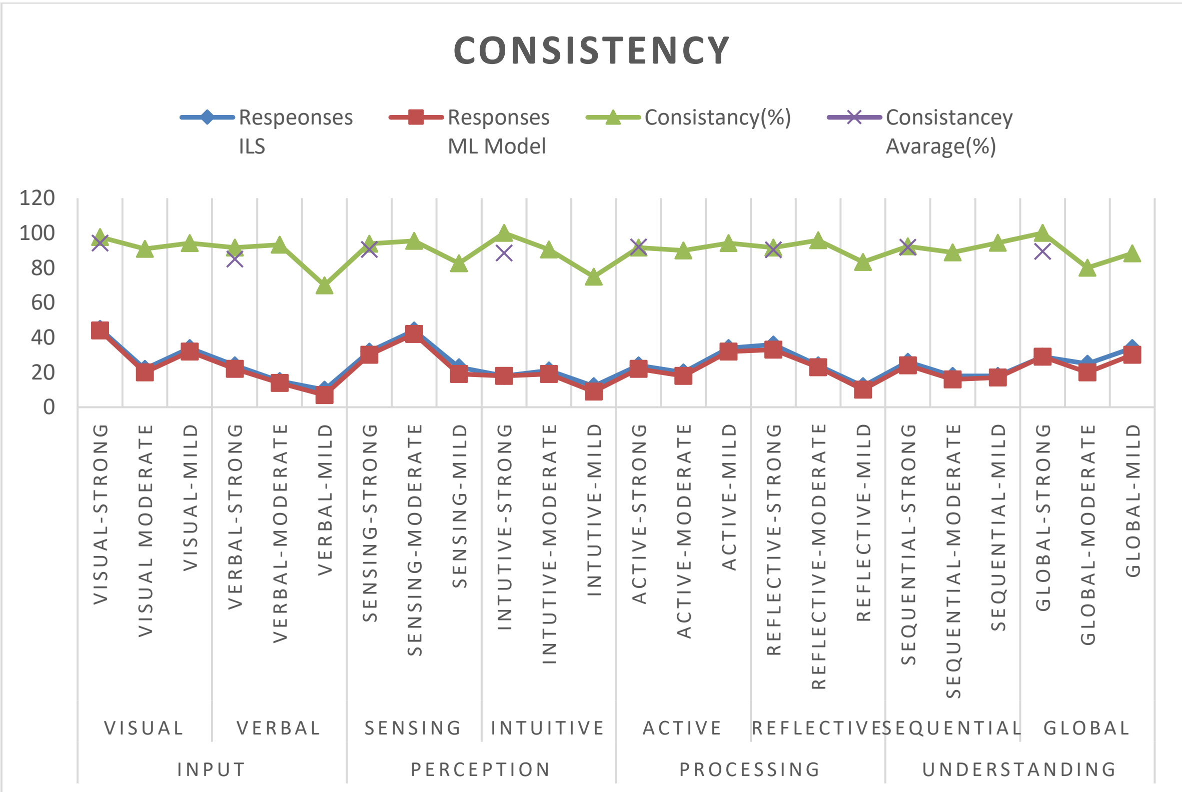


Fig. 18. Consistency of ML model

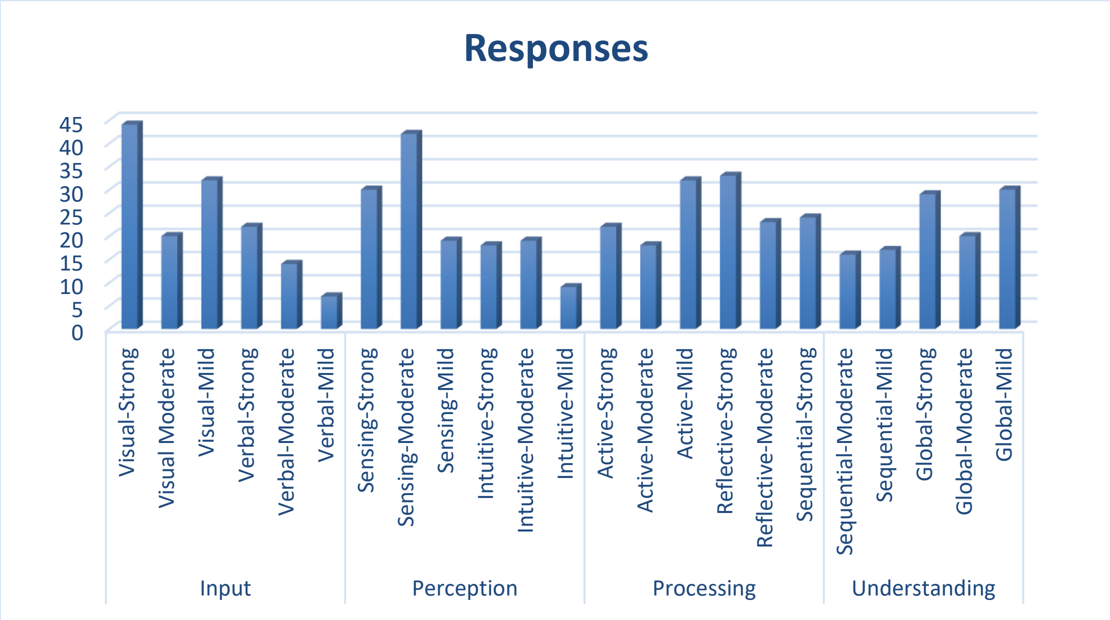


Fig. 19. Summary of the responses

the Learning pattern for the selected data sample can be identified as Global-Mild, Visual- Strong, Sensing- Moderate and Reflective-Strong. This assist teachers in identification of how the course materials should be improved and developed accordingly to cater the requirements of the preferences of the students.

4. CONCLUSION AND FUTURE WORKS

The learners have their own patterns of learning. Hence, Identification of the Learning style of the students plays a pivotal role in designing the learning content in both conventional physical classrooms as well as the online learning environment. The research's main objective was to develop ML models for identifying the learning style in the online learning environment. In the research process, it was identified that Decision Tree Classifier performs with the best accuracy over the other selected classification algorithms. The ML model was validated using K-fold grid validation and MSE, Bias and Variance values were considered for the model evaluation for identifying underfitting and overfitting context.

5. CONTRIBUTION TO THE RESEARCH COMMUNITY

As the research contribution, the experiment has given a new methodology for identification of the learning style of the students in online learning environment according to the FSLSM model. The proposed methodology's significance is that it combined the frequency of access to the course contents and the total time spent on a particular activity. This suggests an improved methodology rather than considering both perspectives separately. In these seven features tailored with the FSLSM model were tested and validated the use of features by considering the relationship to each dimension. Then an application for identification of the learning style considering each four dimensions was designed where teachers can know the learning style of the students and provide the feedback them accordingly. Furthermore, it was designed reusable time tracking and FSLSM questionnaire plugins for the Moodle environment which any of the researcher can use for further research.

6. LIMITATIONS

As for the limitations, it was identified that engaging the students to refer to the course contents is one of the major challenges. And the success of the identifying the learning style of the students using FSLSM indexed learning Style questionnaire depend on responses given by the students with a good understanding of the phenomena of the questions

7. FUTURE WORKS

As the Future works, providing an Artificial Intelligence based recommendation system based on the FSLSM learning style model for improving learning process can be seen as one of the potential research areas. And proposing a methodology for identifying and analyzing the aspects that influence students' attractiveness when reading course materials or attending a presentation. This will be one of the primary areas that will aid in the improvement of course content design.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Galang RE, Hallar AD. The impact and effectiveness of e-learning on teaching and learning, International Journal of Computing Sciences Research. 2021;5(1): 383–397.
DOI: 10.25147/ijcsr.2017.001.1.47
2. Thwe WP, Kálmán A, Lifelong learning in the educational setting: A Systematic Literature Review, Asia-Pacific Education Researcher; 2023.
DOI: 10.1007/s40299-023-00738-w
3. Kaplan A. Lifelong Learning: Conclusions from A Literature Review; 2016.
Available:www.iojpe.org
4. Koenig E, Guertler K. One Size Does Not Fit All: Individuality and Perceptions of Improvement and Satisfaction Among TE Students, English Teaching and Learning, vol. 2021;45(3):303–324.
DOI: 10.1007/s42321-021-00076-4
5. Rubavathy D. A Study on the preferred learning styles of students. 2013;7(2).
DOI: 10.25215/0702.013
6. Shemshack A, Spector JM. A systematic literature review of personalized learning terms, Smart Learning Environments, Springer. 2020;7(1).
DOI: 10.1186/s40561-020-00140-9
7. Rasheed F, Wahid A. Learning style detection in E-learning systems using machine learning techniques, Expert Syst Appl. 2021;174.
DOI: 10.1016/j.eswa.2021.114774.
8. Musa A. Online Learning and Learning Styles; 2023.
Available:https://www.researchgate.net/publication/326302822
9. Doulik P, Skoda J, Simonova I. Learning styles in the e-learning environment: The approaches and research on longitudinal changes, International Journal of Distance Education Technologies, IGI Global. 2017;15(2):45–61.
DOI: 10.4018/IJDET.2017040104.
10. Sönmez S, Korucuk M. The Effect of E-Learning styles on online learning, Shanlax International Journal of Education. 2023;11:216–226,
DOI: 10.34293/education.v11is1-july.6203
11. Fida A, Ghaffar A. Learning Styles: An overview of the felder-silverman's model and measure; 2015.
12. Tamada MM, Giusti R, De Magalhaes Netto JF. Predicting student performance based on logs in moodle LMS, in Proceedings - Frontiers in Education Conference, FIE, Institute of Electrical and Electronics Engineers Inc; 2021.
DOI: 10.1109/FIE49875.2021.9637274
13. Abdullah M, Aljabri J, Fallatah R, Alqahtani A, Altowirgi R. Learning style classification based on student's behavior in moodle learning management System A novel adaptive e-learning model matching educator-student learning styles based on machine learning View project Wireless Multimedia Sensor Networks View project Learning Style Classification Based on Student's Behavior in Moodle Learning Management System; 2015,
DOI: 10.14738/tmlai.31.868
14. Sianturi ST, Yuhana UL. Student behaviour analysis to detect learning styles using decision tree, Naïve Bayes, And K-Nearest Neighbor method in moodle learning management system, IPTEK The Journal for Technology and Science. 2022; 33(2):94.
DOI: 10.12962/j20882033.v33i2.13665.
15. Ikawati Y, Al Rasyid UH, Winarno I. Student behavior analysis to detect learning styles in moodle learning management system, in IES 2020 - International Electronics Symposium: The Role of Autonomous and Intelligent Systems for Human Life and Comfort, Institute of Electrical and Electronics Engineers Inc. 2020:501–506.
DOI: 10.1109/IES50839.2020.9231567
16. Wanniarachchi WAAM, Premadasa HKS, Supervised Machine learning techniques for students performance prediction-an approach of Feldr Silverman Learning Style (FSLSM), 9th International Conference of Sabaragamuwa University of Sri Lanka – ICSUSL; 2023.

Available:https://www.icsusl.sab.ac.lk/ICSUSL_2023_Book_of_Abstracts.pdf
Accessed: Jan. 20, 2024.

17. Nasir IM et al. Pearson correlation-based feature selection for document classification using balanced training, Sensors (Switzerland). 2020;20(23):1–18.
DOI: 10.3390/s20236793.
18. Benno Haarman BCM, Riemersma-Van der Lek RF, Nolen WA, Mendes R, Drexhage HA, Burger H, Feature-expression heat maps - A new visual method to explore complex associations between two variable sets, J Biomed Inform. 2015;53:156–161.
DOI: 10.1016/j.jbi.2014.10.003
19. Battula K, Research of machine learning algorithms using k-fold cross validation; 2021.
Available:https://www.researchgate.net/publication/349946623
20. Ayeni AW. Empirics of Standard Deviation; 2014.
DOI: 10.13140/2.1.1444.6729.
21. Khan M, Noor S. Performance analysis of regression-machine learning algorithms for predication of runoff time, Agrotechnology. 2019;08(01).
DOI: 10.35248/2168-9881.19.8.187
22. Fahse T, Huber V, van Giffen B, Managing Bias in machine learning projects, in Lecture Notes in Information Systems and Organisation, Springer Science and Business Media Deutschland GmbH. 2021:94–109.
DOI: 10.1007/978-3-030-86797-3_7
23. Erickson BJ, Kitamura F. Performance metrics for machine learning models, Radiology: Artificial Intelligence,. Radiological Society of North America Inc. 2021;3(3).
DOI: 10.1148/ryai.2021200126
24. Marosan Z, Savic N, Klasnja-Milicevic A, Ivanovic M, Vesin B. Students' Perceptions of ILS as a learning-style-identification tool in e-learning environments," Sustainability (Switzerland),. 2022;14(8).
DOI: 10.3390/su14084426.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/112361