# Identifying Driver behaviors in learning style classification Using Machine Learning Technique

**AHMAD AL-QEREM[1], MOHMMAD AL_QEREM[2], ISSAM JEBREEN[3]**

[1]Department of Computer Science, Faculty of information technology, Zarqa University, Zarqa, Jordan
ahmad_qerm@zu.edu.jo
[2]Al-Ahliyya Amman University Faculty of Arts & Sciences Humanities & Social Sciences Department
M.alqerem@ammanu.edu.jo
[3]Department of Software Engineering, Faculty of information technology, Zarqa University, Zarqa, Jordan
ijebreen@zu.edu.jo

## ABSTRACT

The most educational institutions commonly and successfully apply learning Management Systems (LMSs) as WebCTs, Blackboards and Moodles. While they are focused on helping teachers to create and run online classes, the individual difference of learners is generally not taken into consideration. However, there is a specific need for learners to have such as preceding experience, cognitive attributes, motivation and learning types. More recent consideration has been focused on aspects such as styles of learner's personality, its impact on learner's personality, and how learning structures could reflect these human characteristics.

These research studies are based on educational theories that argue that courses that fit students' particular characteristics facilitate learning for them and therefore increase learning progress. The focus of this study is to extend the LMS so that learning styles are adapted to the two popular learning models Feld-Silverman and Kolb.

**Key words:** learning style, Machine learning, Driver behaviors, E-learning.

## 1. INTRODUCTION

Nowadays, massive amounts of data is generated and available from various fields and sources. Due to the size of the data and its varied format and sources, it is difficult for humans to sift through it and extract value. For this reason, machine learning is becoming the new method through which data is analyzed for value. Machine learning is being used to analyze data, create rules and find hidden patterns and relations that could not have been found by humans. It is being used in a wide range of fields, such as search engines, image processing, fraud detection, etc.

This research aims to expand MMS to adapt to the two popular models Feld-Silverman and Kolb's Learning Styles by incorporating learning styles. When studying classification styles, each style of learning is determined by the behavior (features) of the student who classifies the student to a specific style of learning. The question is how to select the most appropriate grouping of learning styles by the best subset of behaviors (drivers). In many conducts the student style is selected, our objective in this study is to

model the relationship between them so as to determine the difference between them.

Through extension of LMSs with adaptively, a support learning environment is implemented to help both parties (teachers and learners). In such an adaptation of LMS, teachers still take advantage of the LMSs properties as well as learners continue benefit of courses adaptation. This study provides a new way to enhance an LMS by focus on leaners need and characteristics, paid attention to them directly, and offer learners with adaptive courses in which adaptation is always enhanced and improved to a learner' needs.

Machine learning works by 'learning' from historical data, to build a mathematical model that is able to predict a future state. There are two types of 'learning': supervised and unsupervised. Unsupervised learning works by grouping alike objects together through common behavior or characteristics. Supervised learning requires that the output (or the variable which we are trying to predict) is previously known. It then creates a model that uses the different attributes of the data available to predict the class attribute. The purpose of this paper is to predict learning style using different machine learning methods and feature engineering techniques. It works on different data sets are used (two data sets) and applies three supervised machine learning methods to each in order to predict learning style. It also shows how applying feature engineering techniques can lead to even better results. The importance of predicting learning style is that it helps institutes run intervention programs to increase student retention rates.

## 2. RELATED WORK

The prediction of learning style is a common field of research in academic literature. Most of the research followed a similar pattern: different machine learning methods are used for predictive analysis and their results are compared to find which method proved most accurate.

One study was done through three machine learning methods and then the results of these conducted methods have been compared to evaluate which method was most effective in prediction.

Similarly, the J48 classifier proved the most effective at

predicting student dropout at the "Eindhoven University of Technology". The conducted of ADT model (decision tree model) produced highest accurate results when applied to a dataset about university sample from three different universities from India, and the GATREE (genetically evolved decision trees) model proved best when used in a study at the Hellenic Open University to analyze distance education data.

Another study done at the University of Jordan explored applying ensemble methods along with different machine learning methods, and compared their results. The study also built these methods with and without behavioral features. Decision trees models are given the best accurate results, and building the model with behavioral features boosted these results even further.

Other study at Minho University was conducted, it showed that applying decision trees model, neural networks model, random forest model and support vector machine models to a dataset which included the results of past exams gave better predictive performance than on a data set which did not include past results.

This paper provides and adds to the current studies of learning style a comparison of method selection and feature engineering techniques that applied through machine learning in order to show how they offer an effective approach to enhance prediction of learning style.

## 3. METHODOLOGY

This paper uses supervised machine learning methods to predict learning style. Prediction analysis works by learning from previously seen and known data to build a model, which is then applied to similar unseen data, to predict a certain feature of the data (output class.)

The past known data is called the testing data set. A model is built based on the testing data set, and is applied to a previously unseen part of the data, called the training data set. The predicted outcome from the test set is compared with the actual values to measure the accuracy and effectiveness of the model.

This study conducted two different datasets. The first dataset is from the University of Jordan. It is made up of 480 student records and 17 attributes relating to student data. The second one is from the Minho University, Portugal. It has 395 student records and 31 attributes. Each of the data sets where split into 75% training sets and 25% testing sets.

The three predictive machine learning methods used to predict learning style in this paper are linear regression, decision tress and Naïve Bayes Classifier. Moreover, regarding a finite set of relations between variables of data (dependent, and independent = "predictive outcome") a Linear regression method has been conducted in order to creates a straight line that best fits the data and generalizes these relations [1].

A graph structure of decision trees, where a node has been created once a potential decision generated, representing as a graph of tree-like [2].

Naïve Bayes classification relies on Bayes' Theorem:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Here, the independent events reflected by A and B variables and, the equation P(A|B) is the probability of A occurring given that B has already occurred. The phrase "naïve" is crated based on the independence hypothesis [3].

In addition to these methods, this paper applies feature engineering techniques to the data sets before feeding them into the models. Feature engineering means selecting certain features, removing unnecessary features or creating new features to feed into the model in order to improve its results and speed up the algorithm.

One way to selecting features is based on their relevance and impact on the dependent variable. This relevance can be calculated based on the correlation between the feature(s) and the dependent variable. Another way is to create a custom variable – as a combination of several independent features, for example – so it can be used better in prediction. Needless to say, feature engineering requires domain knowledge so the right variables are selected, modified or created.

## 4. EVALUATION METHODS

To assess the model effectiveness, the predicted outcome from the test set is compared with the actual values to measure how accurate the predictions are. One way of doing this is by using the confusion matrix [4] shown in table 1 below:

**Table 1:** Confusion matrix

|  | Predicted as true | Predicted as true |
|---|---|---|
| **Actually True** | True positive | False Negative |
| **Actually False** | False positive | True Negative |

We can get three different evaluation criteria from a confusion matrix. The first is accuracy, which is the ratio of correct predictions. It is calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

However, accuracy is not always a good evaluation method. This is because it overshadows the minority cases in the data. For instead, a dataset of 100 records of the students, 90 of which are students who succeeded and 10 who failed, has an accuracy of 90%. Meaning that predicting that every new student will 'succeed' is 90% accurate, which is incorrect.

Therefore, two other evaluation criteria are used, precision and recall. They are calculated as follows:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Precision represents the successful positive predictions of the model and recall represents successful negative predictions. The main idea is that in order to have best predictive model a combination of both precision and recall must occurs. This combination is represented by F-measure value that is used for final comparisons as evaluation criteria F-measure value calculation is shown below:

$$F\text{-}measure = 2.\frac{Precision.Recall}{Precision+Recall}$$

## 5. DATA SETS

Technological enhanced learning, in recent years, focused on models of learning styles in order to make e-learning systems more adaptable. With the inclusion of not only dimensions of learning styles, but also the different features of these dimensions, the learning styles of the student are represented more precisely and thus the potential of adaptive learning environment is increased. In addition, a thorough investigation into the characteristics of learning styles can also improve pedagogical models to promote more effective and individualized learning.

As mentioned before, Student learning style is showed how student learning techniques and how they prefer to learn, and support that an teacher can effectively teach to conform to individual students, the researcher bring to literature a nice survey as follows: [5 and 6]. Student behavioral characteristics is a one factor of student learning style in which could be obtained and evaluated from a student's learning behaviour. Several models have been deigned and developed to define and measure the learning style, for example (1) Kolb [7], which suggested that students could be categorized into convergence, divergence and accommodation. (2) The student can be identified with alternate treatment, differentiation, analytical talent and spatial skill in the Keefe learning styles test [6]. (3) Model Area & Silverman [8] includes "the intuitive / sensitive, global / sequential, visual / verbal, inductive / deductive and active / reflective group". For example, the sensitive / sequential / verbal / inductive / active is a style of learning. (4) In Stangl, [9], the students differentiated themselves in four styles: acting, listening, reading and seeing. Manish Joshi [10] has found a similarity in the types of learning,

using the Silverman Field model to understand 176 MCA and MCM students at IMCC in Pune. The student distribution in different categories is presented in Table I. As we can see, there are 72.67 percent of students showing both active and reflecting styles; 59.66 percent showed both Sensing and Intuit characteristics. This study compares between kolb's model, and Felder and Silverman model in order to assess students' style of learning.

## 6. KOLB'S LEARNING STYLES

Kolb identified based on the various modes of learning Figure.1, four learning styles: Accommodating Style (Active, Concrete). The student asks the question: what if ...? He primarily learns by "manipulation" and practice. He likes to participate in the design and operating of activities, he uses testing and error rather than logic to solve problems. He tends to rely rather on others' reflections than on his own analysis, accepting risk. He accepts risk taking. Diverging Style (Concrete, Reflective): The student poses the question: why? Why? He has an eye-watching impression. He has the ability for different aspects to perceive objects and problems. He is excellent in innovative activities such as brainstorming, a fertile imagination and varied interests. Type of assimilation (reflective, abstract) The student asks: what? What? He can organize differing information logically. Rather of practical usage he likes to research concepts and hypotheses. Style of convergence (Abstract, Active): The student asks: How? He has the ability to use ideas and theories in practice, solve problems and decide. He tends to solve problems with one solution, however. Finally, it is easier than participation in interpersonal or social issues for him to perform professional tasks.
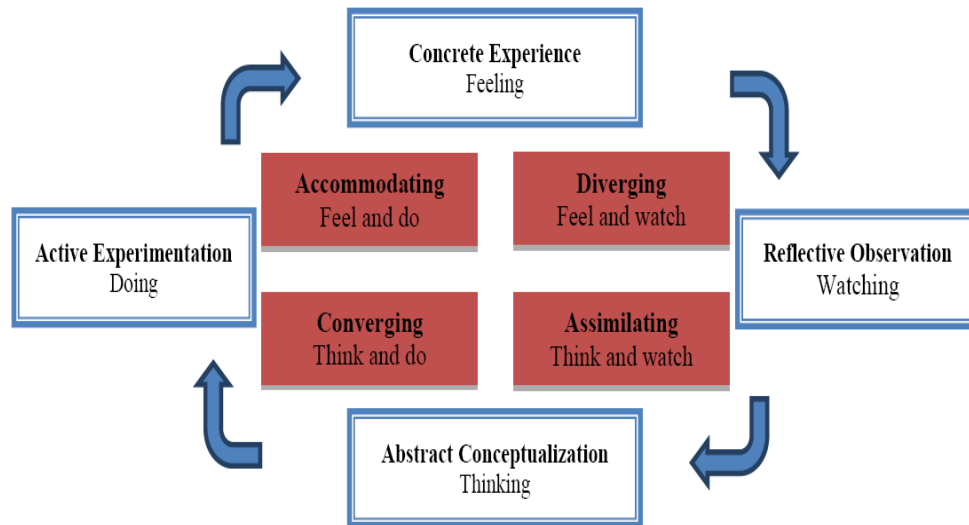


**Figure.1**: Kolb's Learning Styles

## 7. FELDER-SILVERMAN LEARNING STYLES

Felder argues that all variations in the population, which in clude personal preferences for schooling, mental and emoti onal variations, affect an individual's learning style. For hi

m, the student's learning style in the presentation, organiza tion, retrieval and simulation of information is to be taken into account.

- He proposes a model composed of four dimensions with the assumption that the student's learning style may be defined in part by the answers to four questions: How

does the student prefer to process information: "actively" through engagement in physical activity or discussion, or "reflectively" through introspection?

- What type of information does the student preferentially perceive: "sensory" sights, sounds, physical sensations, or "intuitive" memories, ideas, insights?
- Through which modality is sensory information most effectively perceived: "visual" pictures, diagrams, graphs, demonstrations, or "verbal" sounds, written and spoken words and formulas?
- How does the student progress toward understanding: "sequentially" in a logical progression of small incremental steps, or "globally" in large jumps, holistically?

In order to determine the position of a student on a scale, Felder provides the ILS questionnaire (Index of Learning Styles), consisting of 44 closed questions and answered by two (a and b) in 4 groups of 11 questions. The cognizant model of a student consisting of four dimensions (D 1, D2, D3 und D4) is described in each group of questions shown in figure 2. For the following degrees, each dimension ranges from -11 to 11:

- From 1 to 3: uncertain
- From 5 to 7: Moderate
- From 9 to 11: strong

The first dimension D1 varies from active to reflective. Learner with Active characteristic lean towards to better understand information through doing, discussing, applying and explaining to others something active with it. Reflective students prefer to first think calmly. In addition, active students tend to like more than passive students who choose to work alone. Sitting through lectures without physical action, but taking notes is difficult for both styles of learning, but particularly difficult for active learner.

The second dimension, which differs from sensory to intuitive, is D2. Sensors students often like facts about learning and intuitive learners often prefer to discover opportunities and relation. Sensors, too, often prefer to solve issues with well-settled technique and aversion difficulties and astonishments. Sensors are probably to be checked on content not directly addressed in class than initiators. Sensors are often patient with details and well memorized and useful (laboratory) work; intuitors can better understand new concepts and are frequently more comfortable than abstractum-containing sensors.

Sensors are typically cautious and good at memorizing information and working hand-to-do (laboratory). Intuitors

may be more relaxed at understanding new concepts rather than mathematically formulated sensors. Sensors are also more convenient and attentive than intuitors; intuitors are more effective and innovative than sensors. Finally, sensors do not like courses with no obvious relation to reality. There are no "plug-and-chug" courses for intuitive people, which require a lot of memorization and repetitive calculations.

Sensors are typically cautious and good at memorizing information and working hand-to-do (laboratory). Intuitors may be more relaxed at understanding new concepts rather than mathematically formulated sensors. Sensors are also more practical and careful than intuitors; intuitivism are more quickly working and are more innovative than sensors. Finally, sensors don't like races that don't actually relate to realistic. Intuitors don't like "plug-and-chug" races, requiring a great deal of memories and repetitive calculations. The final D4 is the progression, it differs from one sequence to the next. In linear steps, sequential learners tend to understand, each step being logically followed by the previous step.

Global students tend to learn in large jumps, almost randomly absorb material, but suddenly "get" connected. In addition, sequence learners tend to take logical steps in the path to discovery answers; global learners may be able to resolve complex difficulties rapidly or place things collected in new ways, after they have understood the big picture.

Two datasets were used and analyzed in this paper. The first one is from the University of Jordan. It is made up of 480 student records and 17 attributes relating to student data, shown below:

In this data set, the variable we are trying to predict is the 'Class'. It is the grade of the student for the course. It can have three values. A value 'low' means the grade of the student was between 0 – 69. A value 'medium' means the grade of the student was between 70 – 89. A value 'high' means the grade of the student was between 90 – 100. A modification to the class variable was made such that grade 'low' was converted to 0 and grades 'medium' and 'high' were converted to 1. The second data set is from the University of Minho, Portugal. It has 395 student records and 31 attributes, in this data set, the variable we are trying to predict is 'G3'. It can have values between 0 – 20. This variable was also modified such that values 0 – 9 were converted to 0 and values 10 – 20 were converted to 1.
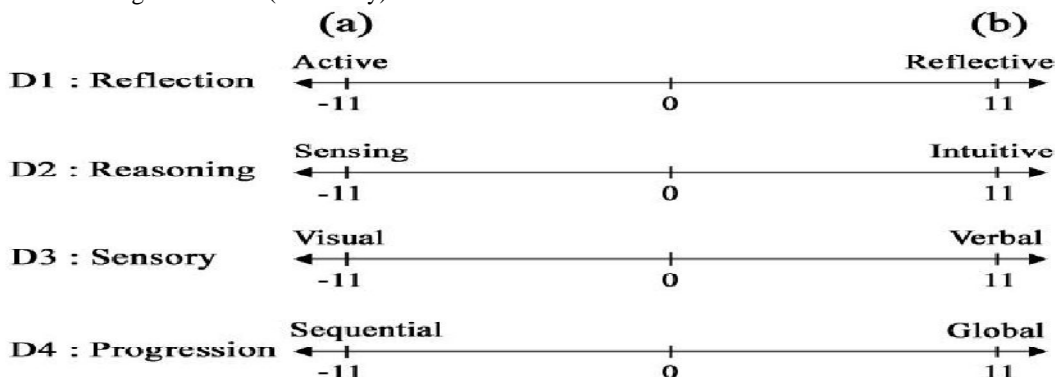


**Figure. 2**: Felder questionnaire of ILS four dimensions

## 8. IMPLEMENTATION

Both datasets were divided to 75% training set and 25% testing set. They were both fed into the three machine learning methods twice, once without modifying the data sets using feature engineering and once after applying feature engineering techniques. The implementation on raw data set 1 Based on the complete data set alone, 353/480 students had "class" (grade) M or H, so the baseline accuracy is 73%. The first model built was linear regression. A confusion matrix was constructed from the results, and it resulted in an improved accuracy of 93%. A decision tree model and naïve Bayes classifier model resulted in 93% and 95% accuracy, respectively. Here, we can already see the positive impact of machine learning algorithms on prediction accuracy. The implementation on raw data set 2 Based on the complete data set alone, 265/395 students had "G3" >=10, so the baseline accuracy is 67%. The first model built was linear regression. A confusion matrix was constructed from the results, and it resulted in an improved accuracy of 74%. A decision tree model and naïve Bayes classifier model resulted in 68% and 73% accuracy, respectively. Here, we can see that applying machine learning algorithms did not enhance prediction accuracy by much.

To explore the effect of applying feature engineering techniques on model performance, two feature engineering techniques were used. The first was variable ranking, to find which features of the data attribute most to the dependent variable. These were found via the conducted of the linear regression methods, which demonstrations the correlation between each independent variable to the dependent variable. The feature creation technique has been applied as second one in which a combination of important features has been applied in order to enhance the structure of decision trees graph and make it more efficient. the implementation on engineered dataset 1 filder solman the models were re-built using only these features. The results for linear regression and Naïve Bayes Classifier had an enhanced accuracy of 95% and 97% respectively.

But for the CART Decision Tree model, the accuracy remained at 93%. So, another technique was applied to the dataset. A custom variable has been created as below:
*customVar= A. StudentAbsences− B. VisitedResources− C. Raisedhands*
The above represents the logic that [a student's success is proportional to "VisitedResources" and "RaisedHands" variables, but inversely proportional to "StudentAbsences" variable]. After building the CART model a third time, the accuracy increased to 96%.

*A. The implementation on engineered dataset 2klob data set*

the models were re-built using only these features. The results for linear regression and Naïve Bayes Classifier had an enhanced accuracy of 77% and 75% respectively.
For the CART Decision Tree model, the accuracy increased to 73%, but further modifications were made to obtain an even

better result. A custom variable has been created as below:
*customVar = A. failures + B. absences − C. studytime*
Therefore, there is inversely proportional of this study time regarding failures and absences. After building the CART model a third time, the accuracy increased to 77%.

## 9. EVALUATION

To show which of the three machine learning methods used was the most effective at predicting learning style, the precision, recall and F-measure values were calculated for each method and summarized in the tables below. The F-measure was the final result used for comparison.

*A. Method comparison for the first raw dataset*
The table 2 below represents method comparison results for the first raw dataset

**Table 2:** comparison results for the first raw dataset

| Method | Accuracy (%) | Precision (%) | Recall (%) | F-measure |
|---|---|---|---|---|
| Linear regression | 93.3 | 93.5 | 97.7 | 0.952 |
| Decision trees | 93.3 | 91.7 | 100 | 0.956 |
| Naïve Bayes classification | 95.8 | 95.6 | 98.9 | 0.972 |

*B. Method comparison for the first engineered dataset*
The table 3 below represents method comparison results for the first engineered dataset

**Table 3:** comparison results for the first engineered dataset

| Method | Accuracy (%) | Precision (%) | Recall (%) | F-measure |
|---|---|---|---|---|
| Linear regression | 95 | 94.6 | 98.9 | 0.97 |
| nDecision trees | 96.7 | 96.7 | 98.9 | 0.98 |
| Naïve Bayes classification | 97.5 | 96.7 | 100 | 0.98 |

The results above show that for both raw and engineered data sets, the use of different machine learning methods did differ in performance much. The Naïve Bayes Classifier produced the highest results overall. However, it can be seen that applying feature engineering techniques did enhance the model performance even further.

*C. Method comparison for the second raw data set*
The table 4 below represents method comparison results for the second raw dataset

**Table 4:** comparison results for the second raw dataset

| Method | Accuracy (%) | Precision (%) | Recall (%) | F-measure |
|---|---|---|---|---|
| Linear regression | 76.4 | 75.5 | 89.4 | 0.826 |
| Decision trees | 68.4 | 71.1 | 89.4 | 0.792 |
| Naïve Bayes classification | 73.5 | 76.3 | 87.9 | 0.816 |

*D.Method comparison for the second engineered data set,*

The table 5 below represents method comparison results for the second engineered dataset.

**Table 5:** comparison results for the second engineered dataset

| Method | Accuracy (%) | Precision (%) | Recall (%) | F-measure |
|---|---|---|---|---|
| Linear regression | 76.9 | 76.1 | 96.0 | 0.87 |
| Decision trees | 78.2 | 76.8 | 96.1 | 0.85 |
| Naïve Bayes classification | 77.2 | 75.9 | 91.9 | 0.84 |

Here, it is seen that the linear regression model produced the best results. The results after applying feature engineering techniques did not matter much for the linear regression and Naïve Bayes models, but increased the Decision tree model's performance considerably.

## 10. DISCUSSION AND CONCLUSION

While the second dataset included additional attributes than the first dataset, the results indicate that first dataset attributes were extra connected to the student success. These illustrations the importance of this data regarding prediction of students' performance.

The results also show that the modification of input data, whether by selecting the most important features of the data or modifying these features themselves, make a big impact on and improve the prediction of learning style.

Finally, the combination of two techniques (feature engineering and method selection) demonstrated the best results even that feature engineering technique was supplementary effective than method selection technique.

## ACKNOWLEDGMENT

## REFERENCES

1. Al Sunbul, H., Silikas, N., & Watts, D. C. (2016). Polymerization shrinkage kinetics and shrinkage-stress in dental resin-composites. Dental Materials, 32(8), 998-1006.
https://doi.org/10.1016/j.dental.2016.05.006
2. Jebreen, I., Al-Qerem, A.: Empirical study of analysts' practices in packaged software implementation at small software enterprises. Int. Arab J. Inf. Technol. 14(4A Special Issue) (2017)
3. Jebreen I., Awad M., Al-Qerem A. (2016) A Propose Model for Shared Understanding of Software Requirements (SUSRs). In: Kim K., Joukov N. (eds) Information Science and Applications (ICISA) 2016. Lecture Notes in Electrical Engineering, vol 376. Springer, Singapore
https://doi.org/10.1007/978-981-10-0557-2_100
4. O. Dahiya and K. Solanki, Comprehensive cognizance of Regression Test Case Prioritization Techniques, International journal of emerging trends in engineering research, Vol. 7 No. 11, pp. 638-646, 2019. https://doi.org/10.30534/ijeter/2019/377112019
5. Aaron Don M., Patrick Bernard T., Arsenic S. Publico, Mharela Angela A. Tan A, "Fuzzy Neural Control System" International Journal of Emerging Trends in Engineering Research, Vol. 7.9 pp. 323-327, 2019. https://doi.org/10.30534/ijeter/2019/15792019
6. O'Keefe, R. M. (1987). What is visual interactive simulation?(and is there a methodology for doing it right?). Institute of Electrical and Electronics Engineers (IEEE).
https://doi.org/10.1145/318371.318635
7. Kolb, D. A. (1984). The process of experiential learning. Experiential learning: experience as the source of learning and development. In : (pp. 20-38). Prentice-Hall, Inc..
8. Silverman, B. G. (1985). Expert intuition and ill-structured problem solving. IEEE Transactions on Engineering Management, (1), 29-33.
9. Monika Hooda, Chhavi Rana, "Learning Analytics Lens: Improving Quality of Higher Education" International Journal of Emerging Trends in Engineering Research, Volume 8. No. 5, pp. 1626- 1646, 2020, https://doi.org/10.30534/ijeter/2020/24852020
10. Joshi, M. (2012). Classification, clustering and intrusion detection system. International Journal of Engineering Research and Applications (IHERA), 2(2), 961-964.