Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 148 (2019) 87–96

Procedia Computer Science

www.elsevier.com/locate/procedia

Second International Conference on Intelligent Computing in Data Sciences (ICDS 2018)

Combining supervised and unsupervised machine learning algorithms to predict the learners' learning styles

Ouafae EL AISSAOUI[a], Yasser EL ALAMI EL MADANI[b], Lahcen OUGHDIR[a] and Youssouf EL ALLIOUI[c]*

[a]*LSI, Sidi Mohamed Ben Abdellah Univesity, Fez, Morocco*
[b]*ENSIAS, Mohammed V Univesity, Rabat, Morocco*
[c]*LS3M, FPK, Hassan First University,B.P. : 145, 25000, Khouribga, Morocco*

Abstract

The implementation of an efficient adaptive e-learning system requires the construction of an effective student model that represents the student's characteristics, among those characteristics, there is the learning style that refers to the way in which a student prefers to learn. Knowing learning styles helps adaptive E-learning systems to improve the learning process by providing customized materials to students. In this work, we have proposed an approach to identify the learning style automatically based on the existing learners' behaviors and using web usage mining techniques and machine learning algorithms. The web usage mining techniques were used to pre-process the log file extracted from the E-learning environment and capture the learners' sequences. The captured learners' sequences were given as an input to the K-modes clustering algorithm to group them into 16 learning style combinations based on the Felder and Silverman learning style model. Then the naive Bayes classifier was used to predict the learning style of a student in real time. To perform our approach, we used a real dataset extracted from an e-learning system's log file, and in order to evaluate the performance of the used classifier, the confusion matrix method was used. The obtained results demonstrate that our approach yields excellent results.

Peer-review under responsibility of the scientific committee of the Second International Conference on Intelligent Computing in Data Sciences (ICDS 2018).

* Corresponding author. Tel.: +212-067-774-7443; fax: +212-052-349-0354.
E-mail address: yelallioui@gmail.com

1877-0509
Peer-review under responsibility of the scientific committee of the Second International Conference on Intelligent Computing in Data Sciences (ICDS 2018).
10.1016/j.procs.2019.01.012

*Keywords:*Web usage mining techniques; machine learning algorithms; K-modes clustering algorithm; Naive Bayes classifier; adaptive e-learning systems; Learning style model;.

1. Introduction

Students have different ways to deal with learning materials and internalize information, these differences determine learning styles. So, every student is characterized by a learning style that determines his/her preferred way to see, process, perceive and understand information. Knowing the learning style helps the E-learning systems to provide personalized contents to students that fit their requirements and enhance the learning process.

Many solutions have been proposed to identify students' learning styles, the traditional one consists in asking students to fill in a questionnaire, this solution has notable drawbacks. Firstly, filling in a questionnaire is a boring task that consumes a lot of time. Secondly, students aren't always conscientious of their learning styles and the importance of the further use of questionnaires, which can lead them to give arbitrary answers, therefore, the results obtained from the questionnaires can be inaccurate and might not reflect the real learning styles of the students. Thirdly, the results obtained from the questionnaires are static, while the learning styles can be changed during the learning process.

To overcome these limitations, many automatic approaches have been proposed, that aim to detect the students' learning style automatically based on the behaviors of the students while they are interacting with the E-learning system. Detecting learning styles automatically has many advantages over the traditional approaches, on one side, it isn't necessary to waste time in using questionnaires since the information can be extracted from the students' interaction with the system. On the other side, the Learning styles detected by the automatic approaches are dynamic and can be changed according to the students' behaviors, while in the traditional approaches the LSs are static.

The automatic detection of learning styles requires the use of a LS model. A LS model aims to classify the students according to the way they prefer to learn with. Many LSMs have been proposed in the literature such as [1][2][3]. .According to the previous researches [4][5], the FSLSM is the most appropriate to implement adaptive e-learning systems .our approach is based on the FSLSM, section 4 presents the reasons why we have chosen it.

In this work, we have proposed an automatic approach to detect the LSs automatically based on the student's behaviors. The students' behaviors were captured from the E-learning platform's log file and then transformed to a set of sequences where each sequence is composed of the learning objects accessed by a student during a session. The learning objects were matched to FSLSM's learning styles combinations. The students' sequences and the matched learning objects were taken as the input to the K-mode algorithm, this unsupervised algorithm was used to map the sequences to the LSCs. the mapped sequences were used then as a training dataset to predict the learning style using Bayes naive classifier.

The rest of this paper is organized as follows. Section 2 describes the learning style model and the algorithms used in our approach. Section 3 introduces a literature review of related work. Section 4 describes the methodology of our approach. The experiments and results are presented in section 5; Finally, Section 6 presents our conclusions and future works.

2. Background

2.1. Learning styles

A learning style refers to the preferential way in which the student perceives, processes, understands and retains information. Due to personality and environmental factors, each student has his own preferred ways of learning, for example, when doing an experiment, some students can understand by following verbal instructions, while others have to physically practice the experiment themselves. These differences in students' learning styles should be considered by the educational systems to enhance the learning process.

Many definitions have been given to the term learning style in the literature. [6] Defined LS as the characteristic strengths and preferences in the ways learners take in and process information. In [7], authors defined it as the "composite of characteristic cognitive, affective, and physiological factors that serve as relatively stable indicators of how a learner perceives, interacts with, and responds to the learning environment. According to[8], LSs refer to "a complex manner in which, and conditions under which, learners most efficiently and most effectively perceive, process, store, and recall what they are attempting to learn.”.

2.2. The Felder and Silverman Learning style Model

A learning styles model enables the classification of learners according to the way they learn. Many learning styles models exist in the literature such as:[9][10][2]. In this work, we have based on the Felder and Silverman model.

The FSLSM uses the notion of dimensions where each dimension contains two opposite categories, and each student has a dominant preference in each dimension's category .the four dimensions of the FSLSM are :processing (active/reflective), perception (sensing/intuitive), input (visual/verbal), understanding (sequential/global). A learning style is defined by combining one category from each dimension.

Active learners tend to retain and understand information best by doing something with the learning material. **Reflective** learners prefer to think about the learning material quietly first. Active learners also prefer to study in a group, while the reflective learners prefer to work individually.

Sensing learners like courses that deal with the real world facts .They tend to be more practical by doing hands-on (laboratory) work and they are more competent than intuitive at memorizing facts. **Intuitive** learners don't like materials that contain a lot of memorization and routine calculations, and they prefer the ones involving abstractions and mathematical formulations. They may be better than sensors at innovating and grasping new concept.

The **visual** learners prefer to see what they learn by using visual representations such as pictures, diagrams, and charts. While the **verbal** learners like information that are explained with words; both written and spoken.

Sequential learners tend to go through the course step by step in a linear way, with each step followed logically by the next one. **Global** learners tend to learn in large jumps, by accessing courses randomly without seeing connections.

2.3. K-modes clustering algorithm

The k-modes is a clustering algorithm that aims to group similar categorical objects into k clusters. This algorithm is an extension of k-means clustering algorithm with three modifications: replacing the Euclidean distance function with the simple matching dissimilarity measure. Instead of means, it uses modes, and instead of updating the centroids it updates modes using a frequency based method [9].

The following steps illustrate how to cluster a categorical data set X into k clusters:

Step 1: Randomly Select k initial modes, one for each cluster.

Step 2: compute the distance from each object to each mode using the dissimilarity Measures, and then associate each object to the cluster whose mode is the nearest to it. This association defines the first k clusters.

Step 3: Update the modes of the newly defined clusters from Step 2 using theorem (1) described below and then retest the dissimilarity between the objects and the updated modes. If an object is found such that its nearest mode belongs to another cluster rather than its current one, reallocate the object to that cluster and update the modes of both clusters.

Step 4: repeat step 3 until no object has moved to another cluster after a full cycle test of the whole data set.

The dissimilarity Measures and the theorem (1) that are used in the previous steps; are described below:

2.3.1. Dissimilarity Measures

To compute the dissimilarity between two objects X and Y described respectively by m categorical attributes values $(x_1, x_2 \ldots x_m)$ and $(y_1, y_2, \ldots, y_m)$, the two following functions can be used:

$$d(X,Y) = \sum_{j=1}^{m} \delta(x_j, y_j) \qquad where\ \ \delta(x_j, y_j) = \begin{cases} 0, x_j = y_j \\ 1, x_j \neq y_j \end{cases} \tag{1}$$

$$d_{x^2}(X,Y) = \sum_{j=1}^{m} \frac{(n_{x_j} + n_{y_j})}{n_{x_j} n_{y_j}} \delta(x_j, y_j) \tag{2}$$

Where n_{x_i} and n_{y_j} are respectively the numbers of objects that have the categories x_j and y_j for attribute j.

2.3.2. How to select a mode for a set

Let $X = \{X_1, X_2, \dots, X_n\}$ a mode of X is a vector $Q = [q_1, q_2, \dots, q_m]$ that minimizes the following equation:

$$D(Q,X) = \sum_{i=1}^{n} d(X_i, Q) \tag{3}$$

Where D can be either equation (1) or equation(2).

Theorem 1: the function $D(Q,X)$ is minimized if $f_r(A_j = q_j \backslash \boldsymbol{X}) \geq f_r(A_j = C_{k,j} \backslash \boldsymbol{X})$ for $q_j \neq C_{k,j}$ for all $j = 1, \dots, m$ where $f_r(A_j = C_{k,j} \backslash \boldsymbol{X}) = \frac{n_{C_{k,j}}}{n}$ is the relative frequency of category $C_{k,j}$ in X.

2.4. Naïve Bayes

The Naive Bayes is a supervised classifier and it is an extension of the Bayes theorem with two simplifications:

The first simplification is to use the conditional independence assumption. That is, each attribute is conditionally independent of every other attribute given a class label C_i .See the following equation:

$$P(C_i \backslash A) = P(C_i) * \frac{P(a_1, a_2, \dots a_m \backslash C_i)}{P(a_1, a_2, a_3, \dots a_m)} = P(C_i) * \frac{P(a_1 \backslash C_i) P(a_2 \backslash C_i) \dots P(a_m \backslash C_i)}{P(a_1, a_2, a_3, \dots a_m)} = P(C_i) * \frac{\prod_{j=1}^{m} P(a_j \backslash C_i)}{P(a_1, a_2, a_3, \dots a_m)} \tag{4}$$

The second simplification is to ignore the denominator $P(a_1, a_2, a_3, \dots a_m)$. Because $P(a_1, a_2, a_3, \dots a_m)$ appears in the denominator of $\boldsymbol{P(C_i \backslash A)}$ for all values of i, removing the denominator will have no impact on the relative probability scores and will simplify calculations.After applying the two simplifications mentioned earlier, we obtain the Naive Bayes classifier which consists in labeling the object A with the class label C_i that maximizes the following equation:

$$P(C_i \backslash A) \sim P(C_i) . \prod_{j=1}^{m} P(a_j \backslash C_i), \ i = 1, 2, \dots, n \tag{5}$$

3. Related works

Enormous solutions have been proposed for automatic detection of learning styles. Many of those solutions are based on data-driven approaches. The data-driven approaches use algorithms from the field of data mining and machine learning to construct a model from the existing students' behaviors and their actual learning styles and then this model is used to identify the learning style for a new student. According to [11]the Bayesian network classifier is one of the most widely adopted classifiers to infer the learning style.

[12] Used a Bayesian network (BN) in order to identify students' learning styles. They identified various behaviors that may be relevant to detect learning styles in a given E-learning system. Then, a BN was trained with data from 50 students, using initial probabilities based on expert knowledge. The trained BN was then evaluated using 27 students. For each student, the

BN provides a probability that the student has a particular learning style preference. As a result, the approach obtained an overall precision of 58% for A/R, 77% for S/I a 63% for S/G (the V/V dimension was not considered).

[13] Addressed this problem by building a BN able to detect the difficulty that learners face during the exploration process; and then providing specific assessment to guide and improve the learner's exploration of the available material.

The decision tree is a classification algorithm that is also frequently used in the field of automatic detection of learning styles. [14] Proposed an approach to automatically detect the learners' learning styles from weblogs of the students using the Data Mining technique and the Decision Trees classifier. Kolb's learning style theory was incorporated to understand e-learners' learning styles on the web. [15] Proposed an approach to classify VARK (Visual, Aural, Read/Write, Kinesthetic) learning styles of learners by using Decision Tree C4.5 algorithm. Data concerning learning styles of learners were collected via a questionnaire responded by 1,205 students. The collected data were then classified using the Decision Tree C4.5 algorithm.

Neural networks are also commonly used in the automatic detection of learning styles,[16]proposed an approach that uses neural networks to identify and track learners learning styles in order to ensure efficient recommendation of resources. Their work was based on Felder-Silverman's dimensions.[17] Proposed the use of a feed-forward ANN (a 3-layer perceptron) with back propagation under a supervised learning model to identify learning styles. Ten behavior patterns were used as inputs such as what kind of reading material did the student prefer, does the student revise their answers on exams prior to submission and does the student ignore, post or read forums? As output, the neural network produces three values, resenting the learning styles on three of the four learning style dimensions of the FSLSM.

As can be noticed, most of the proposed approaches used FSLSM's dimensions and considered that there are 8 learning styles where each one corresponds to a dimension's preference. In reality, there are sixteen learning style combinations obtained by combining one preference from each dimension. In our proposed approach we have considered sixteen LSC instead of 8.

4. Methodology

In this work we have proposed an approach that uses the web usages mining techniques and machine learning algorithms to construct a model based on existing students behaviors (Fig. 1.). This model can be used to identify the learning styles using input students behaviors data.

The students' behaviors were captured from the web log of the E-learning platform. After preprocessing the log file using the web usages mining techniques, the learners' sequences were extracted to be used as an input to a clustering algorithm. The learners' sequences are composed of the set of LOs accessed by the learner during the learning process, each sequence contains the sequence id, session id, learner id, and the set of learning objects accessed by the learner in a session.

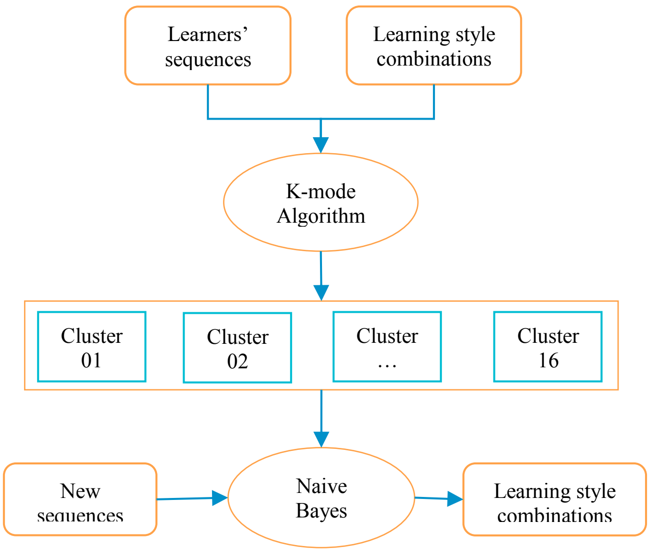


Fig. 1. Our approach

The idea that we have come up with in this work, is to use students' behaviors related to different types of learning objects to determine learning styles. The learning objects were mapped to learning styles using FSLSM, and then a clustering algorithm is used to group the learners' sequences with respect to their learning styles based on the selected Los .for example, if a sequence contains LOs such as: video, image, chart then this sequence will be classified as having a visual preference. In this work, as Fig. 1 illustrates, two machines learning algorithms have been used:

- K-modes algorithm: is used to cluster the learners' sequences into FSLSM's learning styles combinations.
- Naïve Bayes: is used to predict the learning style combination for a new sequence.

The next subsections are organized as follows: first, we will represent how to match learning objects with learning style combinations, then we will describe how the clustering and classification algorithms were used in our approach.

4.1. Matching learning objects to learning styles combinations

The automatic detection of learning styles requires the use of a LSM, in our approach we have based on the FSLSM for many reasons. Firstly; this model uses the concept of dimensions and describes thoroughly the learning style preferences. This description mentions the types of learning objects that can be included in each learning style preference, and this is an interesting characteristic to our work, because knowing the learning style preferences of learning objects helps us to determine the learning style of learners' sequences. Secondly; according to our approach a LS can be changed over time, and then updated using a classifier technique. The FSLSM provides this possibility by considering LS as tendencies and students can act in a non-deterministic way as pointed by[18].

Thirdly , the FSLSM is the most used in adaptive e-learning systems and the most appropriate to implement them as mentioned by [5].

The FSLSM is composed of 4 dimensions, where each dimension contains two opposite learning styles preferences, and each learner has a dominant preference in each dimension. Thus, to identify the learner' learning style; we have to determine a combination composed of one learning styles preferences from each dimension. As a result we will obtain sixteen combinations:

Learning Styles Combinations (LSCs) = {(A,S,Vi,G), (A,S,Vi,Seq), (R,S,Vi,G), (A,S,Ve,Seq), (A,S,Ve,G), (R,S,Ve,Seq), (R,S,Ve,G), (A,I,Ve,G), (A,I,Vi,Seq), (A,I,Vi,G), (R,I,Vi,Seq), (R,I,Vi,G), (R,S,Vi,Seq), (A,I,Ve,Seq), (R,I,Ve,Seq), (R,I,Ve,G)}.

The above 16 learning style combinations have to be matched to learning objects, because identifying the learning style combination of a sequence depends on knowing the learning styles of its learning objects. Based on the matching table presented in our previous works [19] [20] [21]; we have obtained the following table, where the mapped learning objects are considered as feature values of the K-modes clustering algorithm.

Table 1. Matching learning objects to LSCs

Cluster ID	Cluster meaning	Videos	PPTs	Demo	Exercise	Assignments	PDFs	Announcements	References	Examples	Practical Material	Forum	Topic list	Images	Charts	Email	Sequential
C01	R,I,Ve,G	yes	yes	no	yes	yes	yes	yes	yes	No	no	yes	yes	no	no	yes	no
C02	A-I-Ve-G	yes	yes	yes	yes	yes	yes	yes	yes	No	no	yes	yes	no	no	yes	no
C03	R-S-Vl-G	yes	yes	no	yes	yes	yes	yes	yes	yes	yes	no	yes	no	no	yes	no
C04	A-S-Ve-G	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	yes	no	no	yes	no
C05	R-I-Vi-G	yes	yes	no	yes	yes	yes	yes	yes	No	no	yes	yes	yes	yes	no	no
C06	A-I-Vi-G	yes	yes	yes	yes	yes	yes	no	yes	No	no	yes	yes	yes	yes	no	no
C07	R-S-Vi-G	yes	yes	no	yes	yes	yes	yes	yes	yes	yes	no	yes	yes	yes	no	no
C08	A-S-Vi-G	yes	yes	yes	yes	yes	yes	no	yes	yes	yes	no	yes	yes	yes	no	no
C09	R-I-Vel-Seq	yes	yes	no	yes	yes	yes	yes	yes	No	no	yes	yes	no	no	yes	yes
C10	A-I-Ve-Seq	yes	yes	yes	yes	yes	yes	yes	yes	No	no	yes	yes	no	no	yes	yes
C11	R-S-Ve-Seq	yes	yes	no	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	yes	yes
C12	A-S-Ve-Seq	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	yes	yes

C13	R-I-Vi-Seq	yes	yes	no	yes	yes	yes	yes	yes	No	no	yes	yes	yes	yes	no	yes
C14	A-I-Vi-Seq	yes	yes	yes	yes	yes	yes	no	yes	No	no	yes	yes	yes	yes	no	yes
C15	R-S-Vi-Seq	yes	yes	no	yes	yes	yes	yes	yes	Yes	yes	no	no	yes	yes	no	yes
C16	A-Se-Vi-Seq	yes	yes	yes	yes	yes	yes	no	yes	Yes	yes	no	no	yes	yes	no	yes

4.2. The K-mode clustering algorithm

The first step of our approach aims to group the learners' sequences into 16 clusters, and then labeling the resulted clusters with the appropriate LSCs. We have used the k-modes clustering algorithm because of its ability to deal with categorical objects, since the learners' sequences which were considered as the input to that algorithm have categorical attribute values.

After extracting the learners' sequences from the log file using web usage mining techniques, we can use them as an input to the K-modes by turning them into a matrix with M rows corresponding to M sequences and sixteen columns to store the attribute values where the attributes of each sequence correspond to the sixteen Los presented in the previous mapping table (Table 1).

Let $S = \{S_1, .. S_i, \dots, S_n\}$ be a set of n sequences (categorical objects), each sequence S_i is defined by 16 attribute values $(s_{i1}, .. s_{ij}, \dots, s_{i16})$.

And let $A_1, .. A_j, \dots, A_{16}$ be the 16 attributes describing the n sequences where:
$\boldsymbol{A}_1$ = video, $\boldsymbol{A}_2$ = PPTs, $\boldsymbol{A}_3$ = demo, $\boldsymbol{A}_{4,}$ = Exercise, $\boldsymbol{A}_5$ = Assignments, $\boldsymbol{A}_6$ = PDFs , $\boldsymbol{A}_7$ = Announcements, $\boldsymbol{A}_8$ = References, $\boldsymbol{A}_9$ = Examples, $\boldsymbol{A}_{10}$ = Practical Material, $\boldsymbol{A}_{11}$ = Forum, $\boldsymbol{A}_{12}$ = Topic list, $\boldsymbol{A}_{13}$ = Images, $\boldsymbol{A}_{14}$ = Charts, $\boldsymbol{A}_{15}$ = Email, $\boldsymbol{A}_{16}$ = Sequential.

For each sequence $\boldsymbol{S}_i$, Each attribute $\boldsymbol{A}_\mathrm{j}$ has two possible categorical attribute values: {yes or no, yes if the j^{th} learning object exists in the sequence, No if it doesn't exist.

Therefore, each sequence is presented as $(s_{i1}, .. s_{ij} \dots s_{i16})$ where $s_{\mathrm{ij}} = \{yes\ or\ no\}$

To perform the k-modes algorithm we used the R software framework for statistical analysis and graphics, and in order to write and execute the R code easily, we used the $RStudio$ as a graphical user interface. The dataset employed in our approach was extracted from the E-learning platform's log file[†] of Sup' Management Group[‡]. This dataset records 1235 learners' sequences. The following table displays the results obtained after running the k-modes with the R.

Table 2. Result of the K-modes algorithm

Cluster ID	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16
Number of sequences in each cluster	58	73	147	49	82	56	187	59	35	32	119	50	46	96	37	109
Total	1235															

The obtained clusters were labeled with the LSCs based on the minimum distances between the clusters' Modes and the LSCs' vectors presented in table I. the distances are computed using the dissimilarity measure described in section 1 .

4.3. Naive Bayes Algorithm

After applying the K-modes algorithm and labeling the sequences with the LSCs, the labeled sequences were used as a training dataset to train the classification algorithm, and then using it to predict the LSC for a new sequence.

[†] http://www.supmanagement.ma/fc/login/index.php
[‡]http://www.supmanagement.ma/fc/

In our work we have applied the Naive Bayes classifier for many reasons. First of all, because it's one of the most efficient machine leaning algorithms, it learns fast and predicts equally so, and it doesn't require lots of storage. A very important characteristic of Naïve Bayes for our work is that it is a probabilistic classification method, therefore, in our approach we consider that the learner's LSC is not deterministic and not stationary since it can be changed during the learning process in an unexpected way, thus we can measure the LSC for a given learner after each iteration using a probabilistic method.

Given a sequence S_{i} defined with 16 attributes $(A_1, .. A_j, \ldots, A_{16})$, as it is described in the previous subsection above, each attribute $\boldsymbol{A}_j$ has two possible attribute values $a_{\mathrm{j}} = \{yes\ or\ no\}$depending on the existence of the j^{th} in the sequence.

Given a set of classified labels $\mathrm{C} = \{C_1, c_2, \ldots C_{16}\}$ where C_{i} corresponds to one of the sixteen LSCs.
According to the Naive Bayes Classifier, for each new sequence S_{i}, we will assign the classifier label C_{i} that maximizes the following equation:

$$P(C_i). \prod_{j=1}^{16} P(a_j \backslash C_i), i = 1, 2, \ldots, 16 \tag{6}$$

To predict the LSC for new sequences, we used the Byes naïve in R with the package e1071, and in order to evaluate the performance of the used classifier, we did an experiment using the confusion matrix in R with caret. The following section describes the experiment steps and the obtained results.

5. Experiment and results

5.1. Confusion metrics for classification problems

To evaluate the performance of the naive Bayes classifier we have used the confusion matrix method because of its ability to deal with the multi-class Classification problems .the confusion matrix is a specific table layout that enables for each class label to visualize the number of true positives, true negatives, false positives and the false negatives. The following metrics can be computed based on the confusion matrix:

$P_i = \dfrac{TP_i}{TP_i + FP_i}$: *Precision (Positive Predictive value) is the fraction of true positive instances among the predicted positive instances.*

$NPV_i = \dfrac{TN_i}{TN_i + FN_i}$: *Negative Predictive Value (NPV) is the fraction of true negative instances among the predicted negative instances.*

$R_i = \dfrac{TP_i}{TP_i + FN_i}$: *Recall (Sensitivity) is the proportion of positive instances that are correctly classified as positive.*

$S_i = \dfrac{TN_i}{TN_i + FP_i}$: *Specificity (True negative rate) is the proportion of negative instances that are correctly classified as negative.*

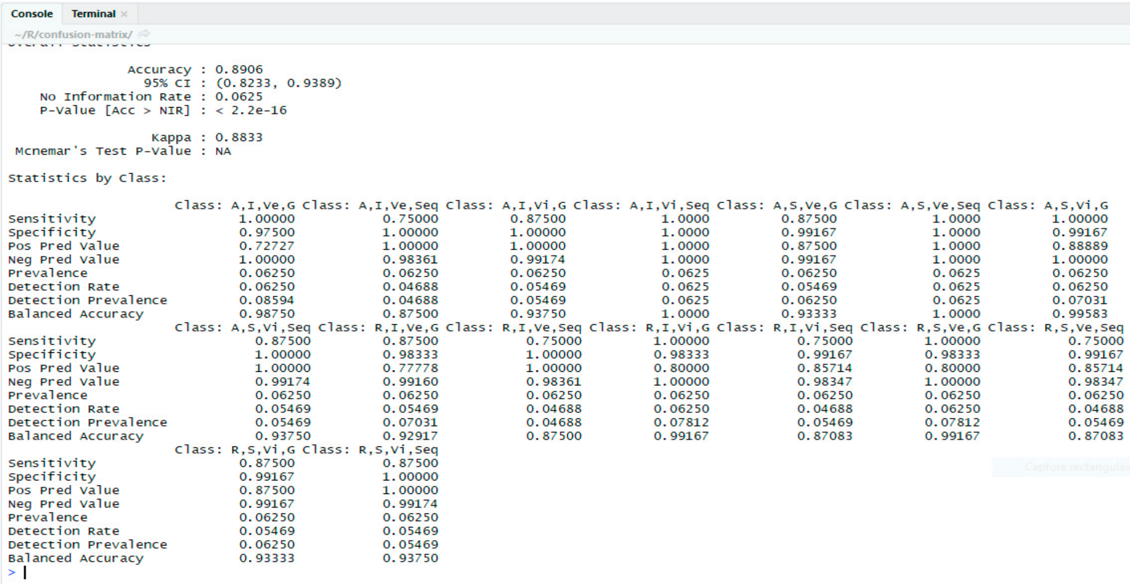
$A = \dfrac{TTP_{all}}{total\ number\ of\ classifications}$: *Accuracy, defines the rate at which a model has classified the records correctly.*

5.2. Experiment

The validation set approach involves randomly dividing the available set of observations into two parts, training set and testing set or hold-out set. The model is fitted on the training set, and the fitted model is used to predict the responses for the observations in the testing set. In our experiment, the available set corresponds to the dataset obtained after performing the K-modes.in this dataset all the sequences are labeled with LSCs, this dataset is split into two parts training and testing with size of 70% and 30% of data respectively. The naive Bayes is fitted on the training set, and then used to predict the responses for the observations in the testing set. The predicted and actual class labels in the testing set are given as parameters to the confusion matrix function .the confusion matrix has been computed using the R with the package caret, the following subsection describes the obtained results.

5.3. Results and discussion

The more the validation metrics are high the more the classifier is good. The captured picture below displays the results obtained after running the confusion matrix using R with the package caret. As can be noticed, all the validation metrics: The Accuracy, the Recall (Sensitivity), the Specificity, the Precision (PPV) and the NPV have high values, so we can say that the classifier used in our approach has well performed.



```
Console   Terminal
~/R/confusion-matrix/

               Accuracy : 0.8906
                 95% CI : (0.8233, 0.9389)
    No Information Rate : 0.0625
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.8833
 Mcnemar's Test P-Value : NA

Statistics by Class:

                     Class: A,I,Ve,G Class: A,I,Ve,Seq Class: A,I,Vi,G Class: A,I,Vi,Seq Class: A,S,Ve,G Class: A,S,Ve,Seq Class: A,S,Vi,G
Sensitivity                  1.00000           0.75000         0.87500            1.0000         0.87500            1.0000         1.00000
Specificity                  0.97500           1.00000         1.00000            1.0000         0.99167            1.0000         0.99167
Pos Pred Value               0.72727           1.00000         1.00000            1.0000         0.87500            1.0000         0.88889
Neg Pred Value               1.00000           0.98361         0.99174            1.0000         0.99167            1.0000         1.00000
Prevalence                   0.06250           0.06250         0.06250            0.0625         0.06250            0.0625         0.06250
Detection Rate               0.06250           0.04688         0.05469            0.0625         0.05469            0.0625         0.06250
Detection Prevalence         0.08594           0.04688         0.05469            0.0625         0.06250            0.0625         0.07031
Balanced Accuracy            0.98750           0.87500         0.93750            1.0000         0.93333            1.0000         0.99583
                     Class: A,S,Vi,Seq Class: R,I,Ve,G Class: R,I,Ve,Seq Class: R,I,Vi,G Class: R,I,Vi,Seq Class: R,S,Ve,G Class: R,S,Ve,Seq
Sensitivity                    0.87500         0.87500           0.75000         1.00000           0.75000         1.00000           0.75000
Specificity                    1.00000         0.98333           1.00000         0.98333           0.99167         0.98333           0.99167
Pos Pred Value                 1.00000         0.77778           1.00000         0.80000           0.85714         0.80000           0.85714
Neg Pred Value                 0.99174         0.99160           0.98361         1.00000           0.98347         1.00000           0.98347
Prevalence                     0.06250         0.06250           0.06250         0.06250           0.06250         0.06250           0.06250
Detection Rate                 0.05469         0.05469           0.04688         0.06250           0.04688         0.06250           0.04688
Detection Prevalence           0.05469         0.07031           0.04688         0.07812           0.05469         0.07812           0.05469
Balanced Accuracy              0.93750         0.92917           0.87500         0.99167           0.87083         0.99167           0.87083
                     Class: R,S,Vi,G Class: R,S,Vi,Seq
Sensitivity                  0.87500           0.87500
Specificity                  0.99167           1.00000
Pos Pred Value               0.87500           1.00000
Neg Pred Value               0.99167           0.99174
Prevalence                   0.06250           0.06250
Detection Rate               0.05469           0.05469
Detection Prevalence         0.06250           0.05469
Balanced Accuracy            0.93333           0.93750
>
```

Fig. 2. The results obtained after applying the confusion matrix to the predicted and actual class labels in the testing set

6. Conclusion

Adaptive E-learning has become a promising solution to enhance the efficiency of online educational systems. A necessary requirement in this solution is the automatic detection of learners' learning style in order to provide well-adapted learning contents.in this context, we have proposed an automatic approach to identify the learners' learning styles using web usage mining techniques and machine learning algorithms.

The E-learning platform’s log file was pre-processed using web usage mining technique in order to extract the learners' sequences. Those sequences were mapped to learning styles combinations using the K-modes clustering algorithm and based on the FSLSM. The labeled sequences were used as training set to train the naive Bayes classifier and predict the learning style combination of a new student. To evaluate the performance of the used classifier we have used the confusion matrix method. The produced results show that our approach performs well. In the future work, we would like to have made a comparative study between the naive Bayes classifier and other machine learning techniques such as the Bayesian network and decision tree.

References

[1] H. K. Laschinger and M. W. Boss, “Learning styles of nursing students and career choices.,” *J. Adv. Nurs.*, vol. 9, no. 4, pp. 375–80, 1984.

[2] R. M. Felder and L. K. Silverman, “LEARNING AND TEACHING STYLES,” *J. Eng. Educ.*, vol. 78, pp. 674–681, 1988.

[3] Biggs, “Study Process Questionnaire Manual. Student Approaches to Learning and Studying.,” *Aust. Educ. Res. Dev.*, p. 53, 1987.

[4] S. Graf and Kinshuk, “Advanced adaptivity in learning management systems by considering learning styles,” in *Proceedings - 2009 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Workshops, WI-IAT Workshops 2009*, 2009, vol. 3, pp. 235–238.

[5] J. Kuljis and F. Liu, “A Comparison of Learning Style Theories on the Suitability for elearning,” *Conference on Web Technologies,*

Applications, and Services. pp. 191–197, 2005.

[6] R. M. Felder, “Matters of style,” *ASEE Prism*, vol. 6, no. 4, pp. 18–23, 1996.

[7] J. W. Keefe, “Learning style: An overview,” *Student Learn. styles Diagnosing Prescr. programs*, vol. 1, pp. 1–17, 1979.

[8] W. B. James and D. L. Gardner, “Learning styles: Implications for distance learning,” *New Dir. Adult Contin. Educ.*, vol. 1995, no. 67, pp. 19–31, 1995.

[9] Z. Huang, “A Fast Clustering Algorithm to Cluster Very Large Categorical Data Sets in Data Mining,” *Res. Issues Data Min. Knowl. Discov.*, pp. 1–8, 1997.

[10] P. Honey and A. Mumford, *The manual of learning styles*, 3rd ed. Maidenhead, United Kingdom: Peter Honey Publications, 1992.

[11] J. Feldman, A. Monteserin, and A. Amandi, “Automatic detection of learning styles: state of the art,” *Artif. Intell. Rev.*, vol. 44, no. 2, pp. 157–186, 2015.

[12] P. García, A. Amandi, S. Schiaffino, and M. Campo, “Evaluating Bayesian networks’ precision for detecting students’ learning styles,” *Comput. Educ.*, vol. 49, no. 3, pp. 794–808, 2007.

[13] A. Bunt and C. Conati, “Probabilistic student modelling to improve exploratory behaviour,” *User Model. User-Adapted Interact.*, vol. 13, no. 3, pp. 269–309, 2003.

[14] A. A. Kalhoro, S. Rajper, and G. A. Mallah, “Detection of E-Learners’ Learning Styles: An Automatic Approach using Decision Tree,” *Int. J. Comput. Sci. Inf. Secur.*, vol. 14, no. 8, pp. 420–425, 2016.

[15] O. Pantho, “Using Decision Tree C4 . 5 Algorithm to Predict VARK Learning Styles,” vol. 24, no. 2, pp. 58–63, 2016.

[16] B. Hmedna, A. El Mezouary, and O. Baz, “Identifying and tracking learning styles in MOOCs: A neural networks approach,” *Eur. MENA Coop. Adv. Inf. Commun. Technol.*, vol. 520, no. 2, pp. 125–134, 2017.

[17] J. E. Villaverde, D. Godoy, and A. Amandi, “Learning styles’ recognition in e-learning environments with feed-forward neural networks,” *J. Comput. Assist. Learn.*, vol. 22, no. 3, pp. 197–206, 2006.

[18] Kinshuk, T. C. Liu, and S. Graf, “Coping with mismatched courses: Students’ behaviour and performance in courses mismatched to their learning styles,” *Educ. Technol. Res. Dev.*, vol. 57, no. 6, pp. 739–752, 2009.

[19] Y. EL ALLIOUI, “Advanced Prediction of Learner’s Profile based on Felder Silverman Learning Styles using Web Usage Mining approach and Fuzzy C-Means Algorithm,” *Int. J. Comput. Aided Eng. Technol.*, vol. “in press,” no. forthcoming.

[20] O. El Aissaoui, Y. El Madani El Alami, L. Oughdir, and Y. El Allioui, “Integrating web usage mining for an automatic learner profile detection: A learning styles-based approach,” in *2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, 2018, pp. 1–6.

[21] O. El Aissaoui, Y. El Alami El Madani, L. Oughdir, and Y. El Allioui, “A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments,” *Educ. Inf. Technol.*, vol. in press, pp. 1–17, Sep. 2018.