İSTANBUL UNIVERSITY PRESS

Acta Infologica
**acin** 2024, 8 (2), 260–273
DOI: 10.26650/acin.1567382

Research Article

# A New Approach to Grouping Learners Based on Behavioral Engagement in CSCL Environments

Souhila Zerdoudi[1], Houda Tadjer[1], Yacine Lafifi[1], Zohra Mehenaoui[1]

[1]Computer Science Department, LabSTIC Laboratory, University 8 Mai 1945, Guelma, Algeria

**Corresponding author :** Yacine Lafifi
**E-mail :** lafifi.yacine@univ-guelma.dz

## ABSTRACT

In Computer-Supported Collaborative Learning (CSCL) environments, forming a group is essential for the success of the learning process. Furthermore, several studies on forming groups in CSCL environments have been conducted recently to form ones that promote learners' engagement and collaboration among their members. Forming group-based approaches requires data on learners' actions (or traces) during the learning process. In this study, behavioral traces of learners are used to form groups. In other words, we used a clustering algorithm based on learners' behavioral engagement to form homogeneous groups of learners. The learners must have different levels of engagement within each group to enhance their engagement and cognitive levels. The basis of the proposed grouping algorithm is a set of indicators of learners' engagement. Furthermore, the proposed approach is based on an artificial intelligence algorithm, the k-means clustering method, which is used to find the maximum possibilities for the best clusters. Then, another algorithm is applied to obtain groups of learners with different levels of behavioral engagement. The validation of the proposed approach on a dataset containing behavioral traces from 100 learners was encouraging and promoting.

**Keywords:** CSCL, Group formation, K-means, Learning analytics, Learner engagement

| | |
|---|---|
| **Submitted** | : 15.10.2024 |
| **Revision Requested** | : 13.12.2024 |
| **Last Revision Received** | : 17.12.2024 |
| **Accepted** | : 23.12.2024 |

## 1. INTRODUCTION

Numerous studies have been conducted on strategies and tactics for grouping learners. The criteria for group formation were cognitive profile, learning style, and soft skills. The aim is to obtain homogeneous or heterogeneous groups according to the objective of the grouping process. All actions performed by learners are saved and input to grouping algorithms. Thus, what standards and components should be applied to determine the best method for assigning students to groups in CSCL (Computer-Supported Collaborative Learning) environments?

Our challenge is to combine learning groups with practical and academic skills. We must determine each learner's behavioral engagement level to build a group of learners. First, behavioral engagement refers to learners' participation in extracurricular activities, attendance, and disciplinary actions are all critical indicators of behavioral engagement (O'Donnell & Reschly, 2020). Fredricks et al. (2004) claimed that behavioral engagement involves student behavior and participation in educational and school-related activities. According to (Reschly et al., 2017), this term encompasses good classroom conduct and engagement in extracurricular activities.

In general, learners' engagement is characterized by a qualitative value, such as being very engaged or quietly engaging. Therefore, measuring this skill is a challenge in online learning environments that support collaboration (like CSCL environments). Traces left by learners during their learning process can be used as data to calculate the value of the learner's engagement degree. However, many research questions can be addressed: How can the degree of engagement of learners be calculated, and what indicators are used to calculate this value? Then, after determining this value, how do we group learners into groups with learners having different engagement values? The aim is to obtain groups of learners who can benefit from each other on one side and all the obtained groups that are homogeneous on the other side. In other words, the objective is to form homogenous groups with heterogeneous learners within the groups.

This study simulates a learner grouping system based on behavioral indicators to place students into diverse learning groups.

This paper comprises five sections and is organized as follows. Section 2 discusses related works on learners' engagement and grouping. The third section presents the proposed approach for grouping learners based on behavioral engagement. The simulation study will be presented and discussed in section four. Finally, the last section of this paper presents the conclusions of this study and future research directions.

## 2. RELATED WORKS

Computer Supported Collaborative Learning (CSCL) aims to maximize student learning achievement by promoting student interaction and collaboration through the use of technology in collaborative learning environments (Stahl et al., 2006; Scheuer et al., 2010). To optimize the collaboration in CSCL contexts, the authors in (Kirschner et al., 2002) and (Isotani et al., 2013) outline certain factors that need to be considered to gain a deeper understanding of (a) How to assign students to groups and (b) How to enhance student participation and communication in group work.

Some researchers have studied the tools and techniques used to improve learners' engagement, such as the "Students' Engagement Scale" tool. Sun and Rueda (2012) developed the student engagement scale to assess student engagement in online learning environments. Furthermore, many techniques are used to improve learners' engagement. Learning analytics are among them. Karaoglan et al. (2022) adopted personalized meta-cognitive feedback support based on learning analytics in online learning for learners' recommendations and guidance to improve student engagement.

Some researchers have developed a Social Learning Analytics "SLA toolkit, " which combines social network analysis with lexical analysis to produce information on student forum participation. Hence, the toolbox promotes the behavioral engagement of learners (Chen et al., 2018; Ouyang et al., 2021). Additionally, visualization of the topic network enhances all students' perspective expressions, indicating that demonstrating students' interest in topics can increase cognitive engagement in terms of students' levels of knowledge sharing, construction, and creation (Ouyang et al., 2021).

Learner engagement can be categorized into several types: cognitive, emotional, social, behavioral, etc. The literature reports several ways to measure students' cognitive engagement in classroom and online learning environments. The Online Student Engagement (OSE) Questionnaire is one such instrument. The questions in the Feedback activity were primarily open-ended and aligned with the OSE questionnaire and the "Sloan" instrument for measuring student satisfaction in online courses (Rajabalee & Santally, 2021). Other researchers have used different concepts and technologies for curriculum design, teaching methods, assessments, and the range of academic support that should be included in open-access-enabling courses (Atherton et al., 2017).

Collaboration among learners can improve their engagement level. Furthermore, learning in groups can improve the outcomes of learners and their profiles. In the education field, numerous approaches have been proposed to address the issue of group formation. Bekele (2005) stated that the formation of learning groups is centered around three main points:

- Group size based on learning objectives
- Dividing learners into groups
- Heterogeneity within groups.

While group homogeneity is a requirement that guarantees the group's productivity (Anzieu & Martin, 1971), it is a criterion that ensures the diversity of the members' viewpoints, ideas, and personalities (Bekele, 2005).

The formation group process can be performed manually or automatically (Matazi et al., 2014). Forming a group manually involves either self-selection or instructor selection (Resta & Laferrière, 2007; Srba & Bielikova, 2014; Ounnas et al., 2007). In the self-selection approach, members have the right to choose the most appropriate group. The second approach is managed by the instructor, who decides which members will comprise the group (Abnar et al., 2012; Srba & Bielikova, 2014). This type of selection guarantees better results in balanced grouping; however, it is a reasonably complex process when many members are grouped manually (Srba & Bielikova, 2014; Mujkanovic et al., 2012). To form groups, CSCL environments can automatically create groups with or without human intervention (Abnar et al., 2012). Random selection is a way to create groups automatically (Srba & Bielikova, 2014). Other approaches form groups on the basis of the context (Maqtary et al., 2019).

The collaborative learning research community has presented numerous algorithmic approaches to address this difficulty. Probabilistic algorithms, clustering, semantic web, anthologies, and other techniques are examples of such methods (Bouyzem et al., 2021; Combaudon, 2018). Due to its ability to handle several variables and its speed in generating optimal solutions, researchers have recently employed genetic algorithms to execute cluster compositions in CSCL systems (Da Rocha, 2019). However, the number of such studies has been limited. Darwin's theory of evolution provides the foundation for the meta-heuristic theory known as the genetic algorithm (Zheng et al., 2018).

Cole and co-authors (Cole et al., 2021) proposed a machine-learning technique that uses a positive label to predict numerous labels for a given input. The proposed method is predicated on handling labels that are absent in the training set or missing labels. Other researchers have examined the application of multi-label classification models to identify diabetes complications (Zhou et al., 2021).

A technique for automatically creating learning groups for MSCL (Mobile Computer Supported Collaborative Learning) systems was presented by Amara et al. (2016). The k-means algorithm was used for the following three grouping criteria:

1. Leaner's attributes: abilities, age, gender, and religion.
2. Interactions between learners and teachers, including the learner's behavior.
3. Contextual data: learner location and learning time.

Many researchers employ k-means techniques to organize cooperative groups. Each time, a k-means variant is created with a distinct set of parameters based on the situation's specific needs, and the proposed approach is then evaluated. Thus, the CSCL community needs a guide to inform it on how a k-means algorithm can be applied to clustering students and which configurations are appropriate to compose an efficient cluster and improve learning outcomes.

In the Table 1, we present a comparative analysis of some studies on the formation of groups to promote collaborative learning.

**Table 1.** Overview of forming group techniques

| Reference | Techniques | Contributions |
|---|---|---|
| (Christodoulopoulos & Papanikolao, 2007) | Fuzzy C-means for random selection | Homogeneous and heterogeneous groups.<br>Manual interference by the instructor.<br>Equality is provided by the group size. |
| (Abnar et al., 2012) | Genetic algorithm<br>Greedy algorithm | The homogeneous and heterogeneous groupings of mixed groups are formed with equal dimensions.<br>An iterative process that satisfies the teacher-tutor on the grouping. |
| (Jozan & Taghiyareh, 2013) | Genetic algorithm | Inter-homogenous and intra-heterogeneous groups.<br>Three algorithms are compared: random, genetic, and the proposed method. |
| (Amara et al., 2016) | k-means | Three types of grouping criteria :<br>Learners‘ characteristics,<br>Learners' learning behaviors and contextual information. |
| (Cole et al., 2021) | Solve the forming group problem using semi-supervised learning algorithms based on a unique positive label for each element. | Develop a method for performing multi-label learning in environments where only a few labels are available. |
| (Zhou et al., 2021) | Multi-label classification models based on medical data processing algorithms are used to address this problem. | Develop a method for diagnosing diabetic complications based on medical data. |

## 3. A NEW APPROACH TO GROUP FORMATION BASED ON BEHAVIORAL ENGAGEMENT

As mentioned earlier in this paper, there are many forming group techniques. However, few studies have used learners' traces as inputs. In this study, the proposed process relies primarily on the learner's behavior while studying in a distance learning environment.

The proposed approach should meet the following objectives:

1. We construct a model for each learner in the learning environment.
2. A set of indicators is computed using the traces available in the constructed model.
3. Evaluate the engagement level of each learner based on the previously calculated indicators.
4. Group learners according to their behavior, i.e., their level of behavioral commitment in intra-heterogeneous groups.

Our approach comprises three steps (Figure 1):

1. Collecting learners' digital traces.
2. Measuring the learner's engagement level.
3. Automatic grouping of learners.


**Figure 1.** Process of proposed approach.

### 3.1. COLLECTION OF STUDENTS' DIGITAL TRACES

In this module, all learners' traces are collected during their interactions with their peers to measure their behavior engagement. To collect these traces, the learner performs any action from the moment of their first access to the platform until their disconnection. In this study, we separated the collected traces into four categories (Table 2):

- ***Participation traces:*** This category records all actions performed by learners regarding their interaction with other learners when using communication tools (Messaging tool, Forum, etc.) offered by the system. For example, messages sent by the learner, messages answered by the learner, posts (topics, answers), and comments posted by the learner in the system's Forum.
- ***Presence traces*** concern the availability of learners, the number of connections to the system, and so forth.
- ***Effort traces:*** This category contains traces left by learners while completing their learning activities, such as consulting or downloading pedagogical resources.
- ***Meeting deadlines traces:*** consistently meeting deadlines that show strong engagement with academic or professional obligations, indicating effective time management skills and a sense of accountability.

Table 2 determines the learner's behavioral engagement indicators.

**Table 2.** Indicators related to learners' traces

| Traces category | Action |
|---|---|
| *Participation* | Sending emails<br>Consulting emails<br>Responding to electronic messages<br>Creating a new subject in the forum tool<br>Answering some submitted questions |
| *Presence* | System Access<br>Connections to the system |
| *Effort* | Consulting pedagogical resources<br>Downloading pedagogical resources<br>Research tasks on the proposed system |
| *Meeting deadlines* | Submission of assignments/homework within given deadlines |

### 3.2. MEASURING LEARNERS' ENGAGEMENT LEVEL

In this step, we propose a set of indicators to measure each learner's level of engagement based on traces collected according to their behavior. We recall that not all actions have the same weight regarding learners' engagement, so weighing actions according to their importance is essential to obtain a more precise and meaningful measure. For example, viewing an email demonstrates superficial engagement, whereas responding to an email indicates deeper engagement.

In our work, to evaluate learners' engagement levels based on their traces, we create "a points system" where each type of trace is assigned a weight based on its importance. We propose the following weighting:

- Less important trace=1 point.
- Moderately important trace=2 points.
- Important trace=3 points.

For example, viewing an email's content indicates less important engagement than answering an email. Therefore, if we give a weight equal to 1 point for '**Consult emails**, ' we should give a weight equal to 2 points for "**Response to an email**."

To calculate the level of engagement according to each trace category, we used the following formula:

$$\textbf{En\_cat}_k\textbf{(Li)} = \Sigma_{j=1}^{n} P_j * AC_{jk} \ldots\ldots \quad (1)$$

**En_cat$_k$(Li):** The engagement value of learner Li for category k.
**AC$_{jk}$:** Action/trace j for the engagement category k.
**P$_j$:** The weight allotted to each action or trace j corresponds to its level of importance.
**n:** number of actions/traces.

Thus, the behavioral engagement levels $En\_cat_1$, $En\_cat_2$, $En\_cat_3$, and $En\_cat_4$ represent the categories of traces related to *Participation, Presence, Effort, and Meeting deadlines*, respectively. These levels were normalized to a data scale between 0 and 1. This normalization preserves the original value distribution and transforms all values in the interval [0, 1]. For this purpose, min-max normalization is applied (Indira et al., 2019), and the new normalized value **En_cat$_k$(Li)**' is given by the following formula:

$$\textbf{\textit{En\_cat}}_k\textbf{\textit{(Li)'}} = (En\_cat_k(Li) - min(En\_cat_k)) \;/\; (max(En\_cat_k) - min(En\_cat_k)) \quad (2)$$

### 3.3. AUTOMATIC GROUPING OF LEARNERS

In this work, our objective is to form heterogeneous groups of learners based on their behavioral engagement levels in an online learning system. We propose two steps to achieve this: First, we use the k-means technique to form k groups containing the most similar learners (Algorithm 1). The k-means and clustering algorithms, in particular, all have a common goal: grouping similar elements into clusters. These elements can be anything as long as they are encoded in a data matrix. K-means has numerous fields of application; it is notably used in clustering data mining during data exploration to detect similar individuals.

A point (a learner X) is assigned to a cluster according to its distance from a different K centroids. In addition, this point (learner X) is assigned to a cluster if it is closer to its centroid (minimum distance). Finally, the distance between two points in the k-means case is calculated using formula (3).

Generally, depending on the need, other techniques can be used once these populations are detected. Our approach used another method or algorithm to group learners into heterogeneous groups. Second, we use one learner from each group to form new groups with varied engagement levels (see Algorithm 2). From this algorithm, we form heterogeneous groups of learners. These groups are composed according to the learners' activities in the online learning system. In heterogeneous groups, learners are grouped according to differences in levels of engagement. Heterogeneous groups based on abilities positively affect students, regardless of their abilities. However, high-potential students may be dissatisfied because helping low-potential students requires additional effort.

The following figure (c.f. Figure 2) illustrates the learners' dynamic grouping process.

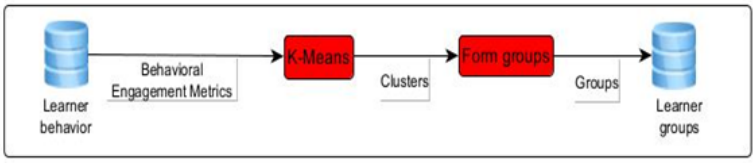


**Figure 2.** Proposed grouping process.

**Step 1: K means clustering**

In this step, we use the K-means algorithm to form groups of learners based on their level of engagement (algorithm 1). K-means clustering is a widely used method for partitioning a set of data points into K clusters, which are measured by the number of connected nearby objects (neighbors) to form a cluster. We use this method to group learners in K clusters based on similarity or distance. The distance between two learners was calculated using the following formula:

$$d(x1, x2) = \sqrt{\Sigma_{j=1}^{n}(x_{1j} - x_{2j})^2} \qquad (3)$$

In this study, the number k of clusters (groups) was equal to the number of categories. In other words, because we have 4 categories, the number of clusters is 4.

**Algorithm 1.** *k-means algorithm*

***Inputs :***

- K: number of groups to form.
- Data matrix: L learners (lines of the data matrix), where each learner Li is represented by K engagement levels (columns of the data matrix) ($En\_cat_{1i}$, $En\_cat_{2i}$, $En\_cat_{3i}$, … $En\_cat_{ki}$)

***Outputs: k initial groups or clusters:*** G1, G2, …, Gk

***Begin***

Randomly select K points (one row of the data matrix). These points are the cluster centers (called centroids).

***Repeat***

Each point (element of the data matrix) is assigned to the group closest to its center.

The center of each cluster is recalculated, and the centroid.

**Until** convergence (or stabilization of the total population inertia)

***End***

**Step 2: Form heterogeneous groups**

Based on the groups obtained in the previous step, we propose to form new groups with different levels of engagement using the following algorithm:

**Algorithm 2.** *New group formation algorithm*

***Inputs : K groups:*** G1, G2, G3, …, Gk

N: Number of learners (***Lij: Learner number i from the group j)***

***Outputs: M** New groups (NG1, NG2, ……, NGM)*

***Begin***

***If N mod K=0, then M=N div K***

***else M=N div K + 1***

***endif***

**For** i:=1 **to** K,

**For** j:=1 **to** M,

Each learner Lij is assigned from group Gi to group NGj.

**If** there are unaffected learners in group Gi, then

Assign these learners to the incomplete groups.

***End For***

***End For***

***End***

### 3.4. ARCHITECTURE OF THE PROPOSED SYSTEM

To validate the proposed approach, we proposed a collaborative learning system composed of groups with different levels of behavioral engagement. Figure 3 shows the architecture of the proposed system, which is composed of the following components:

- Course management module: This module considers all actions related to creating course content (concepts and learning objects).
- The trace management sub-system: It collects the learner's traces and calculates the indicators that help group learning. This sub-system is composed of two modules:
  1. Traces collection module.
  2. Indicators calculating module.
- Grouping module: This module applies the proposed algorithm (described above) for grouping learners into groups with heterogeneous engagement profiles.

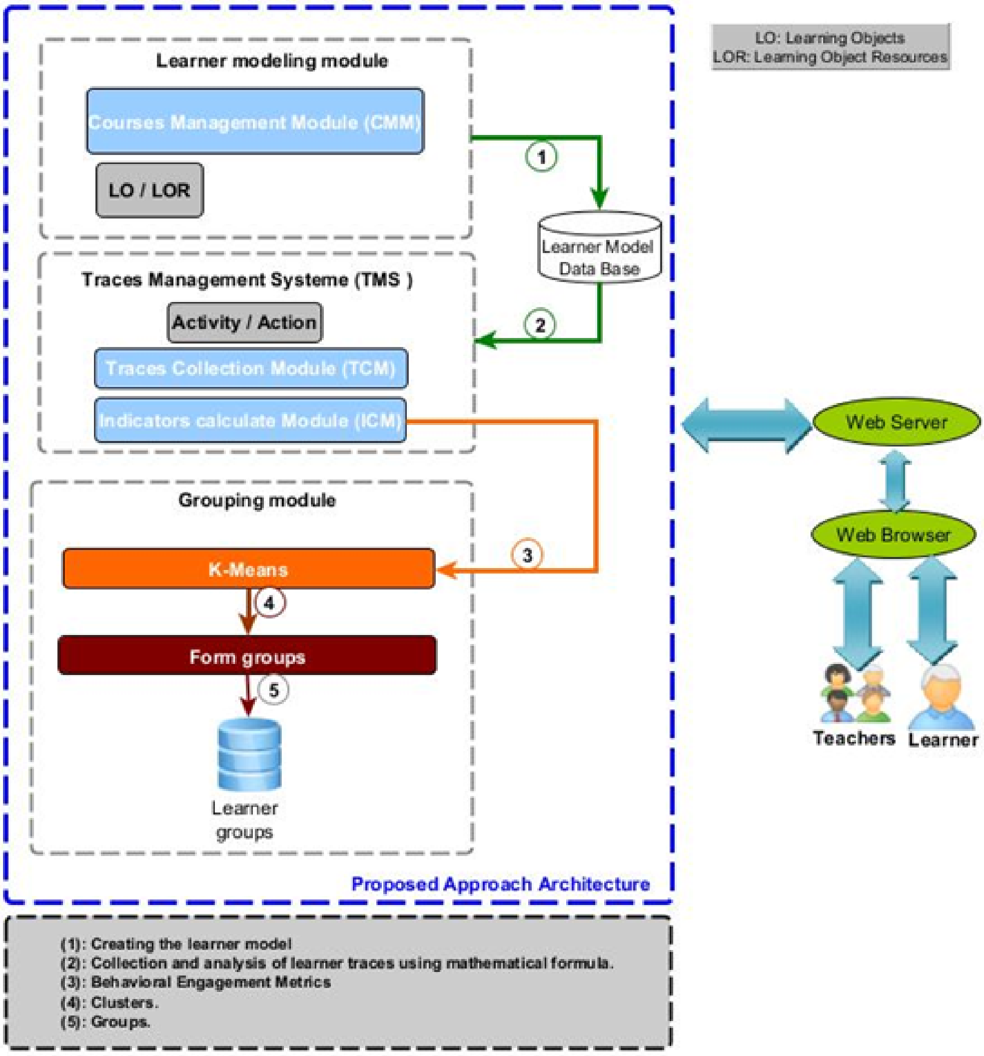


**Figure 3.** Proposed Approach Architecture.

## 4. VALIDATION OF THE PROPOSED APPROACH: RESULTS AND DISCUSSION

To test the proposed approach, we generated a dataset of 100 learners. Therefore, we have 100 learners, indicated by L1 until L100, whose details are given in the Table 3.

**Step 1**

After applying the K-means algorithm to form four (4) clusters, their centers were initially chosen as learners L9, L18, L40, and L22.

**Table 3.** Values of the tested datasets

| Learner | Participation | Presence | Effort | Meeting deadline |
|---|---|---|---|---|
| **L1** | 1.00 | 0.30 | 1.00 | 0.02 |
| **L 2** | 0.06 | 1.00 | 1.00 | 0.97 |
| **L3** | 0.56 | 0.20 | 0.50 | 0.55 |
| **L4** | 0.72 | 0.90 | 1.00 | 0.23 |
| **L5** | 0.89 | 0.90 | 0.50 | 0.58 |
| **L6** | 1.00 | 0.60 | 0.50 | 0.15 |
| **L7** | 0.78 | 0.10 | 0.00 | 0.77 |
| **L8** | 0.50 | 0.30 | 1.00 | 0.79 |
| **L9** | 0.28 | 0.60 | 0.00 | 0.75 |
| **L10** | 0.00 | 0.00 | 0.75 | 0.21 |
| **L11** | 0.94 | 0.70 | 0.00 | 0.65 |
| **L12** | 0.11 | 0.80 | 0.50 | 0.03 |
| **L13** | 0.67 | 1.00 | 0.50 | 0.89 |
| **L14** | 0.78 | 0.90 | 0.00 | 0.26 |
| **L15** | 1.00 | 0.90 | 1.00 | 0.93 |
| **L16** | 0.33 | 0.90 | 0.50 | 0.76 |
| **L17** | 0.83 | 0.70 | 0.50 | 0.59 |
| **L18** | 0.28 | 1.00 | 0.25 | 0.23 |
| **L19** | 0.89 | 0.50 | 0.25 | 0.00 |
| **L20** | 0.72 | 0.90 | 0.75 | 0.49 |
| **L21** | 0.33 | 0.60 | 1.00 | 0.52 |
| **L22** | 1.00 | 0.60 | 0.75 | 0.48 |
| **L23** | 0.22 | 0.80 | 0.75 | 0.32 |
| **L24** | 0.50 | 0.80 | 1.00 | 0.62 |
| **L25** | 0.50 | 0.80 | 0.75 | 0.58 |
| **L26** | 0.83 | 0.70 | 1.00 | 0.55 |
| **L27** | 1.00 | 0.80 | 0.25 | 1.00 |
| **L28** | 0.83 | 0.40 | 0.50 | 0.08 |
| **L29** | 0.89 | 0.60 | 0.50 | 0.55 |
| **L30** | 0.78 | 0.10 | 0.75 | 0.15 |
| **L31** | 0.44 | 0.50 | 0.50 | 0.88 |
| **L32** | 0.11 | 0.40 | 1.00 | 0.35 |
| **L33** | 0.06 | 0.30 | 0.00 | 0.48 |
| **L34** | 1.00 | 0.80 | 1.00 | 0.80 |
| **L35** | 0.83 | 0.40 | 0.75 | 0.28 |
| **L36** | 0.78 | 0.00 | 0.75 | 0.31 |
| **L37** | 0.33 | 0.90 | 0.50 | 0.98 |
| **L38** | 0.33 | 0.90 | 0.00 | 0.62 |
| **L39** | 0.50 | 0.00 | 0.25 | 0.18 |
| **L40** | 0.33 | 0.00 | 0.50 | 0.51 |
| **L41** | 0.12 | 0.55 | 0.4 | 0.33 |
| **L42** | 0.80 | 0.90 | 0.2 | 0.60 |
| **L43** | 0.20 | 0.80 | 0.75 | 0.40 |
| **L44** | 0.60 | 0.40 | 1.00 | 0.90 |
| **L45** | 0.10 | 0.20 | 0.80 | 0.66 |
| **L46** | 0.40 | 0.80 | 0.15 | 0.11 |
| **L47** | 0.95 | 0.12 | 0.67 | 0.32 |
| **L48** | 0.45 | 0.20 | 0.14 | 0.62 |
| **L49** | 0.13 | 0.82 | 0.82 | 1.00 |
| **L50** | 0.56 | 0.22 | 0.34 | 0.18 |
| **L51** | 0.44 | 0.66 | 0.22 | 0.36 |
| **L52** | 0.87 | 0.88 | 0.68 | 0.90 |
| **L53** | 0.99 | 0.17 | 0.55 | 0.30 |
| **L54** | 1.00 | 0.55 | 0.88 | 0.77 |
| **L55** | 0.98 | 1.00 | 0.16 | 0.23 |
| **L56** | 0.64 | 0.66 | 0.48 | 0.66 |
| **L57** | 0.12 | 0.11 | 0.41 | 0.33 |
| **L58** | 0.57 | 0.16 | 0.14 | 0.25 |
| **L59** | 0.46 | 0.14 | 0.28 | 0.64 |
| **L60** | 0.36 | 0.75 | 0.18 | 0.19 |
| **L61** | 0.15 | 0.28 | 0.84 | 0.36 |
| **L62** | 0.32 | 0.19 | 0.68 | 0.12 |
| **L63** | 0.58 | 0.54 | 0.56 | 0.57 |

**Table 3.** Continued

| | | | | |
|---|---|---|---|---|
| **L64** | 0.14 | 0.87 | 0.69 | 0.26 |
| **L65** | 0.55 | 0.57 | 0.49 | 0.83 |
| **L66** | 0.68 | 0.19 | 0.15 | 0.87 |
| **L67** | 0.68 | 0.23 | 0.58 | 0.69 |
| **L68** | 0.94 | 0.88 | 0.46 | 0.78 |
| **L69** | 0.52 | 0.26 | .0.61 | 0.94 |
| **L70** | 0.11 | 0.23 | 0.16 | 0.45 |
| **L71** | 0.28 | 0.16 | 0.58 | 0.17 |
| **L72** | 0.48 | 0.64 | 0.98 | 0.36 |
| **L73** | 0.16 | 0.45 | 0.18 | 0.58 |
| **L74** | 0.45 | 0.88 | 0.26 | 0.17 |
| **L75** | 0.55 | 0.48 | 0.12 | 0.78 |
| **L76** | 0.99 | 0.59 | 0.45 | 0.16 |
| **L77** | 1.00 | 1.00 | 1.00 | 0.68 |
| **L78** | 0.22 | 0.14 | 0.16 | 0.89 |
| **L79** | 0.78 | 0.32 | 0.65 | 0.23 |
| **L80** | 0.50 | 0.30 | 0.45 | 0.59 |
| **L81** | 0.14 | 0.26 | 0.98 | 0.78 |
| **L82** | 0.47 | 0.66 | 0.14 | 0.16 |
| **L83** | 0.78 | 0.69 | 0.57 | 0.78 |
| **L84** | 1.00 | 0.60 | 0.90 | 0.67 |
| **L85** | 0.44 | 0.22 | 0.66 | 0.96 |
| **L86** | 0.66 | 0.35 | 0.68 | 0.45 |
| **L87** | 1.00 | 0.99 | 0.88 | 0.69 |
| **L88** | 0.69 | 0.25 | 0.87 | 0.64 |
| **L89** | 0.32 | 0.66 | 0.17 | 0.98 |
| **L90** | 0.15 | 0.17 | 0.88 | 0.35 |
| **L91** | 0.68 | 0.69 | 0.17 | 0.16 |
| **L92** | 0.45 | 0.80 | 0.90 | 0.74 |
| **L93** | 0.17 | 0.19 | 0.43 | 0.16 |
| **L94** | 0.16 | 0.18 | 0.55 | 0.45 |
| **L95** | 0.22 | 0.80 | 0.67 | 0.78 |
| **L96** | 0.85 | 0.84 | 0.99 | 1.00 |
| **L97** | 1.00 | 0.98 | 0.96 | 0.85 |
| **L98** | 0.65 | 0.79 | 1.00 | 0.98 |
| **L99** | 0.98 | 0.80 | 1.00 | 0.99 |
| **L100** | 0.94 | 0.58 | 0.80 | 1.00 |

Let µ1, µ2, µ3, and µ4 be the centers of gravity of clusters 1, 2, 3, and 4, respectively (See Table 4).

**1st iteration:**

**Table 4.** Gravitational centroids of the first iteration.

| Centers of gravity | The « center » learner | Participation | Presence | Effort | Meeting deadline |
|---|---|---|---|---|---|
| µ1 | L9 | 0.28 | 0.60 | 0.00 | 0.75 |
| µ2 | L18 | 0.28 | 1.00 | 0.25 | 0.23 |
| µ3 | L40 | 0.33 | 0.00 | 0.50 | 0.51 |
| µ4 | L22 | 1.00 | 0.60 | 0.75 | 0.48 |

To calculate the distance, we applied the formula 3. After calculating all the distances between the learners, we find the following clusters were identified:

**Group 1:** L7, L9, L11, L16, L31, L33, L37, L38, L49, L65, L66, L73, L75, L78, L89

**Group 2:** L2, L12, L14, L18, L23, L41, L42, L43, L46, L48, L51, L55, L60, L64, L67, L74 L82, L91

**Group 3:** L3, L8, L10, L30, L32, L36, L39, L40, L45, L50, L57, L58, L59, L61, L62, L69 L70, L71, L80, L81, L85, L90, L93, L94

**Group 4:** L1, L4, L5, L6, L13, L15, L17, L19, L20, L21, L22, L24, L25, L26, L27, L28 L29, L34, L35, L44, L47, L52, L53, L54, L56, L63, L68, L72, L76, L77, L79, L83 L84, L86, L87, L88, L92, L95, L96, L97, L98, L99, L100

**Updating the centers of gravity:** The centers of the formed clusters are represented by averages; thus, we can find new centers of gravity below (See Table 5).

**2nd iteration:**

**Table 5.** Gravitational centroids of the second iteration.

| Centers of gravity | Participation | Presence | Effort | Meeting deadline |
|---|---|---|---|---|
| μ1 | 0.41 | 0.55 | 0.24 | 0.79 |
| μ2 | 0.42 | 0.75 | 0.36 | 0.34 |
| μ3 | 0.35 | 0.18 | 0.60 | 0.44 |
| μ4 | 0.82 | 0.68 | 0.76 | 0.61 |

**The new clusters resulting from the last iteration are as follows:**

**Group 1:** L2, L7, L9, L11, L27, L31, L33, L37, L38, L48, L49, L65, L66, L73, L75, L78 L89, L95

**Group 2:** L12, L14, L16, L18, L19, L23, L28, L41, L42, L43, L46, L51, L55, L60, L64, L70 L74, L82, L91

**Group 3:** L3, L8, L10, L30, L32, L36, L39, L40, L45, L47, L50, L57, L58, L59, L61, L62, L67 L69, L71, L79, L80, L81, L85, L86, L90, L93, L94

**Group 4:** L1, L4, L5, L6, L13, L15, L17, L20, L21, L22, L24, L25, L26, L29, L34, L35, L44, L52, L53, L54, L56, L63, L68, L72, L76, L77, L83, L84, L87, L88, L92, L96, L97. L98, L99, L100

**Step 2**

After obtaining four groups from the previous step, algorithm 2 was used to form new groups. These groups are shown in the Table 6.

**Table 6.** Forming heterogeneous groups

| | 1st Learner | 2nd Learner | 3rd Learner | 4th Learner |
|---|---|---|---|---|
| Group 1 | L2 | L12 | L3 | L1 |
| Group 2 | L7 | L14 | L8 | L4 |
| Group 3 | L9 | L16 | L10 | L5 |
| Group 4 | L11 | L18 | L30 | L6 |
| Group 5 | L27 | L19 | L32 | L13 |
| Group 6 | L31 | L23 | L36 | L15 |
| Group 7 | L33 | L28 | L39 | L17 |
| Group 8 | L37 | L41 | L40 | L20 |
| Group 9 | L38 | L42 | L45 | L21 |
| Group 10 | L48 | L43 | L47 | L22 |
| Group 11 | L49 | L46 | L50 | L24 |
| Group 12 | L65 | L51 | L57 | L25 |
| Group 13 | L66 | L55 | L58 | L26 |
| Group 14 | L73 | L60 | L59 | L29 |
| Group 15 | L75 | L64 | L61 | L34 |
| Group 16 | L78 | L70 | L62 | L35 |
| Group 17 | L89 | L74 | L67 | L44 |
| Group 18 | L95 | L82 | L69 | L52 |
| Group 19 | L93 | L91 | L71 | L53 |
| Group 20 | L94 | L92 | L79 | L54 |
| Group 21 | L77 | L96 | L80 | L56 |
| Group 22 | L83 | L97 | L81 | L63 |
| Group 23 | L84 | L98 | L85 | L68 |
| Group 24 | L87 | L99 | L86 | L72 |
| Group 25 | L88 | L100 | L90 | L76 |

### 4.1. DEGREE OF HETEROGENEITY

As mentioned in this paper, each learner is represented by his engagement level, which is described as a quadruplet (Participation, effort, presence, meeting deadline). We propose assigning a color to each category according to its value to obtain groups of learners with heterogeneous behavioral engagement levels. The colors of each category are as follows:

If En_cat$_i$ ∈ [0, 0.2[, then En_cat$_i$(color)="Red"
If En_cat$_i$ ∈ [0.2, 0.4[, then En_cat$_i$(color)="Orange"
If En_cat$_i$ ∈ [0.4, 0.6[, then En_cat$_i$(color)="Blue"
If En_cat$_i$ ∈ [0.6, 0.8[, then En_cat$_i$(color)="Purpil"
If En_cat$_i$ ∈ [0.8, 1], then En_cat$_i$(color)="Green"

For each learner, we have four colors (CC1, CC2, CC3, CC4), and CCi is the color of the category i.
Therefore, in a group composed of four learners L1, L2, L3, and L4, we obtain the following:
Li(CCi1, CCi2, CCi3, CCi4), where CCij is the color of category j of learner i.
This study aimed to form groups with members with heterogeneous levels of each of the four categories (in other words, having different colors). To this end, we propose the following algorithm to calculate the degree of heterogeneity of each group.

**Algorithm calcul_degree_group_het (K: integer);**
**Input: En_cat (Color) of all learners:(CCi1, CCi2, CCi3, CCi4)**
**Output: Degree_group_het$_k$** //the heterogeneity degree of group K
**Begin**
**number_color$_i$ (CCi1, CCi2, CCi3, CCi4)=j** // j ϵ **[1, 4]**
**If j=1 then het_$cat_i$=0**
Else het_$cat_i$=((j-1) +1/(5-j))/ 4
Degree_group_het$_k$=$\Sigma_{i=1}^{4}$(het_cat$_{ik}$)/4
**End.**

**Example:**
We assume we have a group G1 composed of learners represented by their engagement levels.
L1(Red, Green, Green, Green)
L2(Red, Green, Blue, Red)
L3(Blue, Orange, Blue, Blue)
L4(Green, Blue, Green, Red)
Therefore, the *Degree_group_het* of this group is 0.54

- number_color$_1$(Red, Red, Blue, Green)=3 and het_cat$_1$=2, 5/4=0.62
- number_color$_2$(Green, Green, Orange, Blue)=3 and het_cat$_2$=2, 5/4=0.62
- number_color$_3$(Green, Blue, Blue, Green)=2 and het_cat$_3$=1.33/4=0.33
- number_color$_4$(Green, Red, Blue, Red)=3 and het_cat$_4$=2, 5/4=0.62

*Degree_group_het=0.62+0.62+0.33+0.62=2.19/4=0.54*

### 4.2. DEGREE OF HETEROGENEITY IN THE PROPOSED APPROACH FOR FORMING GROUPS BASED ON LEARNERS' BEHAVIORAL ENGAGEMENT

We obtained a set of groups by applying the two steps of the proposed method. To determine whether the formed groups contain learners with heterogeneous engagement levels, we calculated the degree of heterogeneity for these groups. As a result, Table 7 summarizes the degree of heterogeneity among the formed groups using our proposed approach.

**Table 7.** Degree of heterogeneity in intra-group.

| Groups | D_H_Sub-Profile 1 | D_H_Sub-Profile 2 | D_H_Sub-Profile 3 | D_H_Sub-Profile 4 | A_D_H |
|---|---|---|---|---|---|
| **Group1** | (Red, Red, Blue, Green)=0.62 | (Green, Green, Orange, Blue)=0.62 | (Green, Blue, Blue, Green)=0.33 | (Green, Red, Blue, Red)=0.62 | 0.54 |
| **Group2** | (Purpil, Purpil, Blue, purpil)=0.33 | (Red, Green, Orange, Green)=0.62 | (Red, Red, Green, Green)=0.62 | (Purpil, Orange, Purpil, Orange)=0.62 | 0.54 |
| **Group3** | (Orange, Orange, Red, Green)=0.62 | (Purpil, Green, Red, Green)=0.62 | (Red, Blue, Purpil, Blue)=0.62 | (Purpil, Purpil, Orange, Blue)=0.62 | 0.62 |
| **Group4** | (Green, Orange, Green, Green)=0.33 | (Purpil, Green, Red, Purpil)=0.62 | (Red, Orange, Purpil, Blue)=1 | (Purpil, Orange, Red, Red)=0.62 | 0.64 |
| **Group5** | (Green, Green, Red, Purpil)=0.62 | (Green, Blue, Blue, Green, )=0.62 | (Orange, Orange, Green, Blue)=0.62 | (Green, Red, Orange, Green)=0.62 | 0.62 |
| **Group6** | (Blue, Orange, Purpil, Green)=1 | (Blue, Green, Red, Green)=0.62 | (Blue, Purpil, Purpil, Green)=0.62 | (Green, Orange, Orange, Green)=0.62 | 0.71 |
| **Group7** | (Red, Green, Blue, Green)=0.62 | (Orange, Blue, Red, Purpil)=1 | (Red, Blue, Orange, Blue)=0.62 | (Blue, Red, Red, Blue)=0.62 | 0.71 |
| **Group8** | (Orange, Red, Orange, Purpil)=0.62 | (Green, Blue, Red, Green)=0.62 | (Blue, Blue, Blue, Purpil)=0.33 | (Green, Orange, Blue, Blue)=0.62 | 0.54 |
| **Group9** | (Orange, Green, Red, Orange)=0.62 | (Green, Green, Orange, Purpil)=0.62 | (Red, Orange, Green, Green)=0.62 | (Purpil, Purpil, Purpil, Blue)=0.33 | 0.54 |
| **Group10** | (Blue, Orange, Green, Green)=0.62 | (Orange, Green, Blue, Purpil)=1 | (Red, Blue, Purpil, Purpil)=0.62 | (Purpil, Blue, Blue, Blue)=0.33 | 0.64 |
| **Group11** | (Red, Blue, Blue, Blue)=0.33 | (Green, Green, Orange, Green)=0.33 | (Green, Red, Orange, Green)=0.62 | (Green, Red, Red, Purpil)=0.62 | 0.47 |
| **Group12** | (Blue Blue, Red, Blue)=0.33 | (Blue, Purpil, Red, Green)=1 | (Blue, Orange, Blue, purpil)=0.62 | (Green, Orange, Orange, Blue)=0.62 | 0.64 |
| **Group13** | (Purpil, Green, Red, Green)=0.62 | (Red, Green, Red, Purpil)=0.62 | (Red, Orange, Red, Green)=0.62 | (Green, Orange, Orange, Blue)=0.62 | 0.62 |
| **Group14** | (Red , Orange, Blue, Green)=1 | (Blue, Purpil, Red, Purpil)=0.62 | (Red, Red, Orange, Blue)=0.62 | (Blue, Red, Purpil, Blue)=0.62 | 0.71 |
| **Group15** | (Blue Red, Red, Green)=0.62 | (Blue, Green, Orange, Green)=0.62 | (Red, Purpil, Green, Green)=0.62 | (Purpil, Orange, Orange, Green)=0.62 | 0.62 |
| **Group16** | (Orange, Red, Orange, Green)=0.62 | (Red, Orange, Red, Blue)=0.62 | (Red, Red, Purpil, Purpil)=0.62 | (Green, Blue, Red, Orange)=1 | 0.71 |
| **Group17** | (OrangeBlue, Purpil, Purpil)=0.62 | (Blue, Green, Orange, Blue)=0.62 | (Red, Orange, Blue, Green)=1 | (Green, Red, Purpil, Green)=0.62 | 0.71 |
| **Group18** | (Orange, Blue, Blue, Green)=0.62 | (Green, Purpil, Orange, Green)=0.62 | (Purpil, Red, Purpil, Purpil)=0.33 | (Purpil, Red, Green, Green)=0.62 | 0.54 |
| **Group19** | (Red, Purpil, Orange, Green)=1 | (Red, Purpil, Red, Red)=0.33 | (Blue, Red, Blue, Blue)=0.33 | (Red, Red, Red, Orange)=0.33 | 0.49 |
| **Group20** | (Red, Blue, Purpil, Green)=1 | (Red, Green, Orange, Blue)=1 | (Blue, Green, Purpil, Green)=0.62 | (Blue, Purpil, Orange, Purpil)=0.62 | 0.81 |
| **Group21** | (Green, Green, Blue, Green)=0.33 | (Green, Green, Orange, Green)=0.33 | (Green, Green, Blue, Red)=0.62 | (Purpil, Green, Blue, Orange)=1 | 0.57 |
| **Group22** | (Blue, Green, Red, Blue)=0.62 | (Purpil, Green, Orange, Blue)=1 | (Red, Green, Green, Blue)=0.62 | (Red, Green, Purpil, Blue)=1 | 0.81 |
| **Group23** | (Green, Purpil, Blue, Green)=0.62 | (Purpil, Purpil, Orange, Green)=0.62 | (Green, Green, Purpil, Blue)=0.62 | (Purpil, Green, Green, Purpil)=0.62 | 0.62 |
| **Group24** | (Green, Green, Purpil, Blue)=0.62 | (Green, Green, Orange, Purpil)=0.62 | (Green, Green, Purpil, Green)=0.33 | (Purpil, Green, Blue, Orange)=1 | 0.64 |
| **Group25** | (Purpil, Green, Red, Green)=0.62 | (Orange, Blue, Red, Blue)=0.62 | (Green, Green, Green, Blue)=0.33 | (Purpil, Green, Orange, Red)=1 | 0.64 |
| **Average_Degree_Heterogeneity_Groups** | | | | | **0.60** |

The greater the degree of heterogeneity, the more distant the learners are (Heterogeneous). Therefore, this degree of heterogeneity provides a better result with low heterogeneity in small groups.

## 4.3. DISCUSSION OF THE OBTAINED RESULTS

An analysis of the results in Table 7 reveals that 18 groups (72%) have learners with at least three different engagement categories. In other words, learners of these groups have complementary behavioral engagement values, which is the main objective of this study. In addition, the five groups were composed of learners having at least two different behavioral engagement values. Finally, only two groups (8%) had learners close to each other regarding behavioral engagement values. In conclusion, these results are encouraging and promotive because more than 70% of the obtained groups contained at least three of four learners with different engagement levels. In other words, learners within the formed groups can share.

During this research, we encountered some obstacles. First, we had to figure out how to measure each learner's behavior during his interactions within the system and with his teammates. The issue lies in measuring qualitative values using learners' actions. These actions were used to calculate learners' engagement levels. Second, another issue is classifying learners' digital traces to measure different engagement levels. Experts in psychology must validate the classification of each action into categories. In addition, associating the weights of each action requires further study using more developed techniques like machine learning techniques, to obtain the adequate weight of each learner's digital trace.

Third, in this study, learners were regrouped into groups according to their engagement levels. The aim of collaboration within the group is not indicated. Therefore, learners can communicate with each other without evaluating their collaboration results, which constitutes a limitation of this research. To address this limitation, we propose creating projects to be carried out during the collaboration process by the members of each formed group. So, we can, in the end, evaluate the regrouping process through the analysis of the results obtained by each group during the realization of the assigned projects.

Finally, we believe that adopting learners' engagement levels as a criterion for forming heterogeneous groups of learners will improve their cognitive levels. Further, researchers in the CSCL field can adapt our approach to form groups of learners to perform collaborative tasks, complete common projects, or resolve common exercises or problems. Furthermore, artificial intelligence techniques, like deep learning techniques, can be used to form groups adaptively in collaborative learning contexts.

## 5. CONCLUSION AND FUTURE WORK

Recently, after the COVID-19 pandemic, we have seen growing interest in adapting online learning platforms. However, after the actual use by students, most users, especially teachers, noticed the students' poor interaction or commitment to the activities requested by them. This situation has enabled researchers to search for methods and techniques to improve students' engagement following distance classes. This involvement has several types: emotional, cognitive, social, and behavioral. In this study, behavioral engagement was used because of its importance in learning.

To improve this type of engagement, we group learners into groups with different levels of behavioral engagement. We must first propose indicators to measure commitment to achieve this goal. These indicators were calculated from traces left by learners during their learning and collaboration processes with other students in the e-learning system.

To answer the questions posed in the introduction section, we can say that in this research, we have proposed a new method of grouping learners based on their levels of behavioral engagement. This level is calculated using four indicators: participation, presence, effort, and meeting deadlines. The latter are calculated according to the learners' traces. In addition, the groups must have complementary levels to enable learners to benefit from each other. To this end, we use the k-means algorithm in the first step of the proposed algorithm. Then, in the second step, a new algorithm was used to obtain groups with heterogeneous or complementary engagement profiles.

We conducted a series of tests on a randomly created dataset to validate the obtained results. The results obtained are considered encouraging and promising. By simulating the grouping method, our system effectively grouped people into different profiles according to complementary engagement levels.

In future work, we plan to test the proposed system in an online learning environment to make learning more user-friendly and effective. We also propose to use other methods for grouping learners based on other clustering algorithms and compare them with the proposed algorithm.

---

**Peer Review:** Externally peer-reviewed.

**Author Contributions:** Conception/Design of Study- S.Z., H.T., Y.L.; Data Acquisition- S.Z., H.T., Y.K.; Data Analysis/Interpretation- S.Z., H.T., Y.L., Z.M.; Drafting Manuscript- S.Z., H.T., Y.L.; Critical Revision of Manuscript- H.T., Y.L., Z.M.; Final Approval and Accountability- S.Z., Y.L., H.T., Z.M.; Technical or Material Support - S.Z., Y.L., H.T., Z.M.; Supervision – H.T., Y.L., Z.M.

**Conflict of Interest:** The authors have no conflict of interest to declare.

**Grant Support:** The authors declared that this study has received no financial support.

---

**ORCID IDs of the authors**

| | |
|---|---|
| Souhila Zerdoudi | 0000-0002-2982-1633 |
| Houda Tadjer | 0000-0001-7624-1343 |
| Yacine Lafifi | 0000-0001-8232-4196 |
| Zohra Mehenaoui | 0000-0002-6732-7839 |

## REFERENCES

Abnar, S., Orooji, F., & Taghiyareh, F. (2012, January). *An evolutionary algorithm for forming mixed groups of learners in web based collaborative learning environments.* In the IEEE international conference on technology enhanced education (ICTEE) (pp. 1-6). IEEE.

Amara, S., Macedo, J., Bendella, F., & Santos, A. (2016). Group formation in mobile computer-supported collaborative learning contexts: A systematic literature review. *Journal of Educational Technology & Society, 19*(2), 258-273.

Anzieu, D., & Martin, J. Y. (1971). *La dynamique des groupes restreints.* Presses univ. de France.

Atherton, M., Shah, M., Vazquez, J., Griffiths, Z., Jackson, B., & Burgess, C. (2017). Using learning analytics to assess student engagement and academic outcomes in open access enabling programmes. Open Learning: *The Journal of Open, Distance and e-Learning,* 32(2), 119-136.

Bekele, R. (2005). *Computer-assisted learner group formation based on personality traits* (Doctoral dissertation, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky).

Bouyzem, M., AL Meriouh, Y., & Moustakim, O. (2021). Le e-learning à l'université Abdelmalek Essaâdi: Une analyse descriptive du point de vue des enseignants. *Revue Economie, Gestion et Société, 1*(32).

Chen, B., Chang, Y. H., Ouyang, F., & Zhou, W. (2018). Fostering student engagement in online discussion through social learning analytics. *The Internet and Higher Education*, 37, 21-30.

Christodoulopoulos, C. E., & Papanikolaou, K. A. (2007, October). *A group formation tool in an e-learning context*. In 19th IEEE international conference on tools with artificial intelligence (ICTAI 2007) (Vol. 2, pp. 117-123). IEEE

Cole, E., Mac Aodha, O., Lorieul, T., Perona, P., Morris, D., & Jojic, N. (2021). *Multi-label learning from single positive labels*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 933-942).

Combaudon, S. (2018). *MySQL 5.7: administración y optimización*. Ediciones Eni.

Da Rocha, H. (2019). *Learn Chart. js: Create interactive visualizations for the web with chart.* js 2. Packt Publishing, Ltd.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Indira, B., Valarmathi, K., & Devaraj, D. (2019). An approach to enhance packet classification performance of software-defined network using deep learning. *Soft Computing*, 23, 8609-8619.

Isotani, S., Mizoguchi, R., Isotani, S., Capeli, O. M., Isotani, N., De Albuquerque, A. R., ... & Jaques, P. (2013). A Semantic Web-based authoring tool to facilitate the planning of collaborative learning scenarios compliant with learning theories. *Computers & Education*, 63, 267-284.b

Jozan, M. M. B., & Taghiyareh, F. (2013). An evolutionary algorithm for homogeneous grouping to enhance web-based collaborative learning. *International Journal of Computer Science Research and Application,* 3(1), 74-85.

Karaoglan Yilmaz, F. G., & Yilmaz, R. (2022). Learning analytics intervention improves students' engagement in online learning. *Technology, Knowledge and Learning*, 27(2), 449-460.

Kirschner, P. A., Jochems, W., Dillenbourg, P., & Kanselaar, G. (2002). *Three worlds of CSCL: Can we support CSCL.* Heerlen: Open University of the Netherlands.

Maqtary, N., Mohsen, A., & Bechkoum, K. (2019). Group formation techniques in computer-supported collaborative learning: A systematic literature review. *Technology, Knowledge and Learning, 24*, 169-190.

Matazi, I., Messoussi, R., & Bennane, A. (2014, May). *The design of an intelligent multi-agent system for supporting collaborative learning.* In the 9th International conference on intelligent systems: Theories and applications (SITA-14) (pp. 1-8). IEEE.

Mujkanovic, A., Lowe, D., Willey, K., & Guetl, C. (2012, June). *Unsupervised learning algorithm for adaptive group formation: Collaborative learning support in remotely accessible laboratories.* In International Conference on Information Society (i-Society 2012) (pp. 50-57). IEEE.

O'Donnell, K., & Reschly, A. L. (2020). Assessment of student engagement. In A. L. Reschly, A. J. Pohl, & S. L. Christenson (Eds.), Student engagement: Effective academic, behavioral, cognitive, and affective interventions at school (pp. 55–76). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-37285-9_3.

Ounnas, A., Davis, H. C., & Millard, D. E. (2007). *Semantic modeling for group formation.* Workshop on Personalisation in E-Learning Environments at Individual and Group Level (PING) at the 11th International Conference on User Modeling UM2007, Corfu, Greece.

Ouyang, F., Chen, S., & Li, X. (2021). Effect of three network visualizations on students' social-cognitive engagement in online discussions. *British Journal of Educational Technology*, 52(6), 2242-2262.

Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy. *Education and Information Technologies*, 26(3), 2623-2656.

Reschly, A. L., Pohl, A., Christenson, S. L., & Appleton, J. J. (2017). Engaging adolescents in secondary schools. *School Mental Health Services for Adolescents*, 45-77.

Resta, P., & Laferrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review*, 19, 65-83.

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning, 5*, 43-102.

Srba, I., & Bielikova, M. (2014). Dynamic group formation as an approach to collaborative learning support. *IEEE Transactions on Learning Technologies, 8*(2), 173-186.

Stahl, G., Koschmann, T., & Suthers, D. (2006). *Aprendizaje Colaborativo apoyado por computador: Una perspectiva histórica.* G. Stahl, T. Koschmann, & D. Suthers, Aprendizaje Colaborativo apoyado por computador: Una perspectiva histórica.(pág. 426). Cambridge: RK Sawyer.

Zheng, Y., Li, C., Liu, S., & Lu, W. (2018). An improved genetic approach for composing optimal collaborative learning groups. *Knowledge-Based Systems*, 139, 214-225.

Zhou, L., Zheng, X., Yang, D., Wang, Y., Bai, X., & Ye, X. (2021). Application of multi-label classification models for the diagnosis of diabetic complications. *BMC Med Inform Decis., 21*(1), 182.

### How cite this article

Zerdoudi, S., Tadjer, H., Lafifi, Y., & Mehenaoui, Z. (2024). A New Approach to Grouping Learners Based on Behavioral Engagement in CSCL Environments. *Acta Infologica, 8*(2), 260-273. https://doi.org/10.26650/acin.1567382