Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 19 No. 12 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v19i12.41181

PAPER

Cooperative Learning Groups: A New Approach Based on Students' Performance Prediction

Zakaria Bousalem[1](✉), **Aimad Qazdar**[2], **Inssaf El Guabassi**[3]

[1]Faculty of Sciences and Technologies, Hassan 1st University, Settat, Morocco

[2]ESTIDMA Laboratory, Ibn Zohr University, Agadir, Morocco

[3]LAROSERI Laboratory, Faculty of Sciences, Chouaib Doukkali University, El Jadida, Morocco

zakaria.bousalem@gmail.com

ABSTRACT

Cooperative learning is a pedagogical approach in which students collaborate in small groups to attain a shared academic objective. In the classroom, cooperative learning aims to enhance learning outcomes by promoting the exchange of information, social, and personal resources among students. Group formation is a critical and complex step that significantly impacts the effectiveness of cooperative learning. In this article, we propose a novel approach for constructing cooperative learning groups that employs machine learning to predict student performance and incorporates the most common grouping strategies to recommend optimal group formation.

KEYWORDS

cooperative learning, technologies enhanced learning, recommendation system, machine learning

1 INTRODUCTION

Cooperative learning is a rapidly developing concept in the field of education. In addition to the traditional modes of learning in-person or virtually, such as receiving instruction from a teacher or individual study, students can also learn from one another. Cooperative learning is characterized by the synergistic collaboration of individual efforts within a group, involving discussions and collective construction of knowledge [21]. This approach shifts the focus from individual or competitive learning to a more collaborative one. Numerous studies have confirmed that cooperation in learning is a critical determinant of success in education [17, 27, 21, 22, 24]. Cooperative learning involves two or more students working together, collectively constructing their knowledge through a continual exchange of their information resources, social resources, and personal resources (see Figure 1). Information resources comprise the knowledge, skills, and experiences that can be taught, transmitted, and shared among students. Social resources refer to an individual's accessibility, sensitivity, and position in a network, thereby facilitating a more effective collaboration with their peers. Personal resources encompass the time and energy that students devote to benefit others.

Bousalem, Z., Qazdar, A., Guabassi, I.E. (2023). Cooperative Learning Groups: A New Approach Based on Students' Performance Prediction. *International Journal of Online and Biomedical Engineering (iJOE)*, 19(12), pp. 34–48. https://doi.org/10.3991/ijoe.v19i12.41181

Article submitted 2023-05-05. Resubmitted 2023-06-08. Final acceptance 2023-06-09. Final version published as submitted by the authors.

The formation of an optimal collaborative learning group enables the maximization of information resource sharing in a socially sensitized environment that promotes sharing and cooperation, without exhausting personal resources in terms of time and energy. Consequently, group formation plays a critical role in the success of cooperative learning.

Establishing a cooperative learning group is a complex and significant step. Selecting suitable individuals for a group contributes to creating environments that foster interaction, exchange, and collaboration [8, 26]. Forming groups without considering the diverse characteristics of individuals (e.g., objectives, academic level, motivation) often leads to issues such as disproportionate participation, demotivation, and resistance to group work in future activities [9].

In this paper, we introduce a novel approach for collaborative learning group formation based on students' performance prediction. This approach combines the most prevalent grouping strategies (student-selected, homogeneous, heterogeneous) with the potential of machine learning to recommend the most optimal group formation. The primary research questions addressed in this article are as follows:

– RQ1: How can machine learning predictions enhance group selection?
– RQ2: How can collaborative learning group (CLG) recommendations be performed based on students' performance predictions?
– RQ3: In what ways can CLG recommendations contribute to improving the quality of education and decision-making?

The remainder of this paper is organized as follows: Section 2 presents the background and related work on the most common grouping patterns. Section 3 describes the proposed approach, provides an overview of the dataset utilized for this study, and illustrates the proposed system, as well as its ability to accurately predict students' performance. The recommendation strategy for cooperative learning groups is discussed in Section 4. Finally, Section 5 concludes the paper and offers directions for future research.

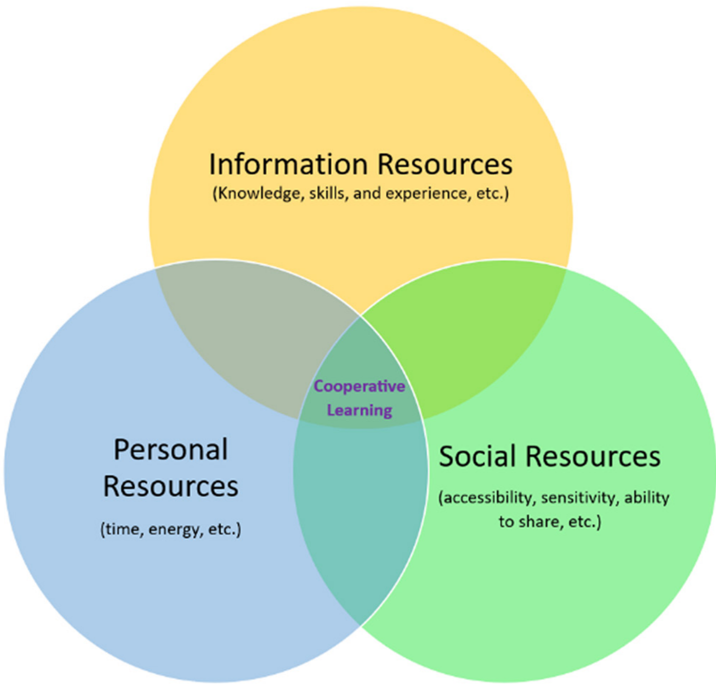


Fig. 1. Interaction of information, social and personal resources in cooperative learning

2 BACKGROUND AND RELATED WORK

2.1 Cooperative Learning Groups (CLG)

Cooperative learning refers to a small group of students working together to achieve a common goal [16]. This approach is a widespread trend in educational strategies, aiming to promote collective intelligence, interactivity, and social development of learners. For a cooperative learning situation to be successful, it must have a commonly accepted goal, upon which the group is rewarded for its efforts. For instance, if a group of learners is assigned to complete a project, but only one student does all the work while the others enjoy a free ride, it is not a cooperative group [25]. Not all group formations of learners working together can be considered cooperative. In fact, some group formations may hinder student learning and create disagreement and discontent, rather than facilitating learning and increasing group synergy [15]. Concerning group work, "group formation is one of the key processes in collaborative learning because having adequate members in the learning groups supports good collaborative interactions among members and is fundamental to ensuring satisfactory learning performance" [6]. The most common grouping patterns include learner-selected groups, teacher-selected groups (either randomly or with a systematic distribution), and automatically selected groups.

2.2 Learner-selected CLG

Learner-selected collaborative learning groups (CLGs) are the most favored group formation by students. In this approach, students choose their collaborators with minimal or no guidance from the teacher. Typically, teachers announce the required number of students per group and allow learners to form groups accordingly. Groups formed by learners tend to exhibit better communication, camaraderie, and enthusiasm about working together. Consequently, learners belonging to learner-selected CLGs often outperform and exhibit less task orientation compared to students in groups formed by other methods [1, 14].

2.3 Teacher-selected CLG

Random CLG: Random collaborative learning groups (CLGs) are formed by assigning students to groups randomly, with consideration given only to the final group size [14]. Random CLGs are the most commonly used grouping strategy in classrooms, as they require little to no preparation or research into students' profiles (ability, academic level, skills, background, etc.) to define their needs for random grouping [20]. Random CLGs exhibit slightly less positive attitudes, lower dynamism, and their performance and outcomes are somewhat inferior compared to those of student-selected groups [20, 5].

Homogeneous CLG: Homogeneous CLGs are formed by teachers where each group member exhibits similar ability, academic level, skills, background, or other characteristics [19]. In most cases, teachers use data from their records of students' academic performance, such as results, marks, grades, and attitudes, to form the groups. Other types of data, such as motivation, interest, and dynamism, can also be used by teachers in this context, although these data are not recorded but can be observed and detected by teachers based on their experiences with students.

Heterogeneous CLG: Heterogeneous CLGs are also formed by teachers; however, unlike homogeneous CLGs, they aim to create well-balanced teamwork comprising students who represent a range of abilities, skills, gender, background, or motivation [28]. Similarly, the data used to place students in heterogeneous groups come from their records of academic performance or profiles. Heterogeneous CLG composition maximizes opportunities for peer support and beneficial mutual exchange [7].

2.4 Automatic-selected CLG

The alternatives for forming CLGs discussed previously have several limitations [18]. For instance, random CLGs may create unbalanced groups, student self-selected groups require good social relations and are not recommended according to the survey conducted by Feichtner [10]. Teacher-selected CLGs cannot handle a large number of students and complex grouping criteria [18]. To address some limitations of these methods, automatic-selected CLGs can be proposed as an alternative solution [18]. Automatic-selected CLGs involve creating groups automatically without any intervention by either the teacher or the student [18]. This method relies on computational techniques, including machine learning, that facilitate automatic selection of student groups based on diverse criteria: preferences, learning styles, academic levels, etc.

As highlighted, this method allows for the creation of homogeneous or heterogeneous groups without any intervention from the learner. However, in our opinion, this may form groups where students will not be confident in their abilities. In this study, we present a mixed approach that leverages the potential of automatic grouping of students based on ML algorithms with the advantages of allowing students to choose their collaborators.

2.5 Related literature

Numerous research studies have been conducted to investigate the impact of cooperative learning groups on enhancing academic performance among learners. According to [22], the use of a homogeneous grouping strategy based on learning styles enables teachers to select at the beginning of the semester the appropriate teaching methodology/strategy and assessment method for in-person classes that suit each student group, rather than treating the classroom as one unit. This approach helps students better achieve the intended learning outcomes of the course [22]. For [28], homogeneous groups are suitable for high-ability students, while heterogeneous groups are appropriate for low-ability students. Van Der Laan Smith and Spindle [28] utilized students' academic performances and perceptions to group students in in-person classes into homogeneous or heterogeneous groups based on their abilities. According to [22], student-selected groups have more trust in each other's abilities and perceive group work as a valuable experience. The findings of [27] show that using heterogeneous groups based on student profiles make students comfortable in their groups and increase content mastery. Lastly, according to our literature review, students prefer working in small groups of 5 to 7 and engaging in smaller assignments, including peer-teaching [17,12,21,3,13].

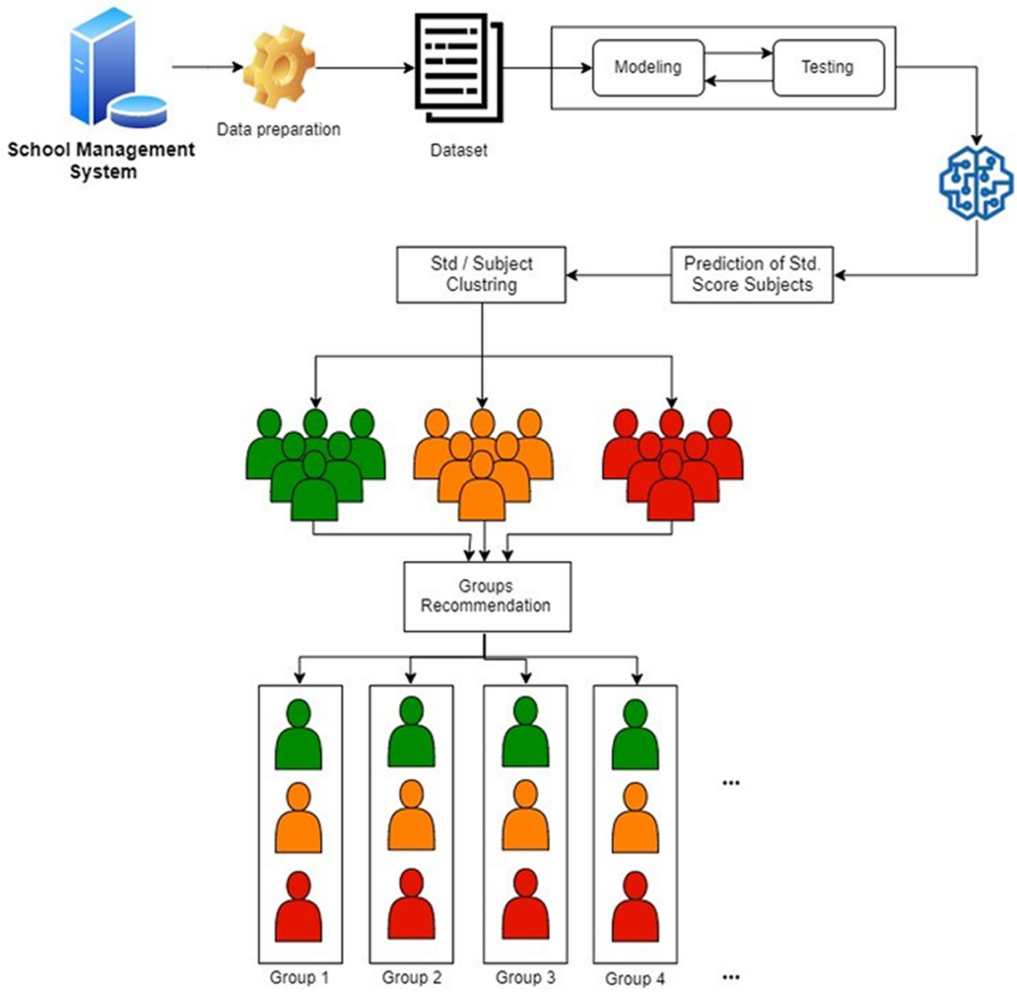


Fig. 2. Recommend cooperative learning groups based on machine learning predictions

Educational Data Mining (EDM), an emergent discipline aiming to develop methodologies for analyzing data from educational settings, is garnering significant attention in our increasingly data-centric world. As a confluence of education, statistics, and informatics, EDM fosters the emergence of sub-domains such as computer-based education, learning analysis, data mining (DM), and machine learning (ML) [29]. The integration of machine learning with EDM, which has seen substantial growth due to the expanding universe of data warehouse sizes, is becoming a necessity and a compelling area of research. This approach has been adopted in diverse fields such as finance, e-commerce, healthcare, tourism, and marketing, including e-learning platforms [30]. Consequently, numerous machine learning algorithms have been leveraged to unearth hidden patterns within educational contexts, underlining the cross-disciplinary impact and relevance of this innovative field [31] [32] [33]. The use of machine learning for group formation has always been of interest to the research community. Numerous studies have been conducted within this domain, utilizing a variety of machine learning algorithms and focusing on specific parameters. [20] created collaborative learning groups of 4 students in virtual learning environments based on collaboration competence levels. To create the collaboration competence levels, two machine learning algorithms for clustering were applied, namely K-Means and Expectation Maximization (EM). K-Means was used in [23] to create heterogeneous groups in in-person classes based on demographic, functional (academic), and personality characteristics. The same algorithm was adopted by [2] to create heterogeneous groups in in-person classes based on learners' preferences. A genetic algorithm was used by [18] to form groups in social network learning environments using academic, cognitive, and social learner characteristics. The datasets used in these works were mostly collected using surveys.

3 PROPOSED APPROACH

In this section, we present an approach based on machine learning algorithms for forming cooperative learning groups. Our approach consists of four main stages:

- Data preparation,
- Model building,
- Student grouping,
- Group recommendation.

Figure 2 provides an overview of the architecture of our proposed approach.

3.1 Data preparation

Fig. 3. Data preparation processing

Data preparation is a crucial step in the field that requires significant effort, but it is essential to contextualize the data to transform it into knowledge and eliminate biases and noise that may impact the quality of data [11]. The data preparation process (see Figure 3) comprises four main operations.

- **Data collection:** This operation involves collecting data on all students enrolled at HBMK High School between 2018 and 2022. The data are exported in a relational format. The personal data fields of students are either eliminated or anonymized and coded to comply with national and international standards for personal data protection.
- **Data filtering:** Data filtering involves selecting Physics stream students' scores, including the Regional score exam that reflects the secondary subjects passed at the end of the first year of Bac, the continuous assessment score, and the scores of eleven subjects of the first semester of the second year of Bac.
- **Data transposition:** In the data collection step, the student information was collected in an inappropriate format for applying machine learning techniques. To address this issue, we used the Talend Open Studio software to transpose the data and aggregate it with the student's code.
- **Dataset:** The dataset comprises actual student data and includes 840 Physics stream students from the 2018–2019, 2019–2020, 2020–2021, and 2021–2022 school years. These are the final available Physics stream data. Table 1 describes the structure of the dataset.

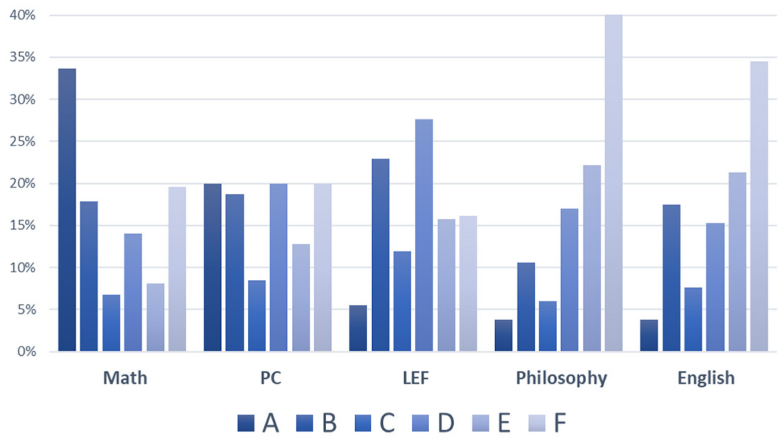


Fig. 4. An overview of student's classification in subjects according to their predictive performance

3.2 Building model

We used various machine learning regression algorithms to train the datasets, including Multiple Linear Regression (MLR), Support Vector Machines (SVM), Decision Tree (DT), Random Forest (RF), Lasso CV (LCV), and Ridge CV (RCV). The goal was to evaluate which algorithm produced a reasonably accurate prediction rate of student performance for effective group recommendation. To achieve this goal, we calculated and interpreted several metrics, including RMSE, EVS, and R^2. As shown in the Table 2, Multiple Linear Regression has the lowest RMSE and highest EVS, R^2, and accuracy compared to other models. Therefore, we can conclude that the Multiple Linear Regression-based model is more efficient than other models in predicting the grades of students.

Table 1. Dataset structure

Attribute	Description	Type
idStd	The identification number of the student	nominal
Sex	The sex of the student	nominal
Age	The age of the student	quantitative
RegScore	The score obtained in the Regional Exam.	quantitative
MathematicsScore1	The first semester's Mathematics mark	quantitative
PhyChemScore 1	The first semester's Chemistry & Physics mark	quantitative
LEF_Score 1	The first semester's Life and Earth Sciences mark	quantitative
PhilosophyScore 1	The first semester's Philosophy mark	quantitative
EnglishScore 1	The first semester's English language mark	quantitative
TranslationScore 1	The first semester's Translation mark	quantitative
PE_Score 1	The first semester's Physical education mark	quantitative
IE_Score 1	The first semester's Islamic Education mark	quantitative
FrenchScore 1	The first semester's French language mark	quantitative
ArabicScore 1	The first semester's Arabic language mark	quantitative

(Continued)

Table 1. Dataset structure *(Continued)*

Attribute	Description	Type
DisciplineScore 1	The first semester's Discipline mark	quantitative
CA1	The first semester's Continuous Assessment score	quantitative
MathematicsScore 2	The Second semester's Mathematics mark	quantitative
PhyChemScore 2	The Second semester's Chemistry & Physics mark	quantitative
LEF_Score 2	The Second semester's Life and Earth Sciences mark	quantitative
PhilosophyScore 2	The Second semester's Philosophy mark	quantitative
EnglishScore 2	The Second semester's English language mark	quantitative
TranslationScore 2	The Second semester's Translation mark	quantitative
PE_Score 2	The Second semester's Physical education mark	quantitative
IE_Score 2	The Second semester's Islamic Education mark	quantitative
FrenchScore 2	The Second semester's French language mark	quantitative
ArabicScore 2	The Second semester's Arabic language mark	quantitative
DisciplineScore 2	The Second semester's Discipline mark	quantitative
CA2	The Second semester's Continuous Assessment score	quantitative
MathematicsScore N	The national exam's Mathematics mark	quantitative
PhyChemScore N	The national exam's Chemistry & Physics mark	quantitative
LEF_Score N	The national exam's Life and Earth Sciences mark	quantitative
PhilosophyScore N	The national exam's Philosophy mark	quantitative
EnglishScore N	The national exam's English language mark	quantitative
NE	The national exam's score	quantitative
GB	Baccalaureate grade	quantitative
Merits	Baccalaureate Merits	nominal

Table 2 represents summary results for the six algorithms used in this research work.

Table 2. Comparative ML algorithm performance

Regressors	RMSE	EVS	R^2	Accuracy (+/–)
Multiple Linear Regression	1.066	0.935	0.935	0.90 (0.09)
Support Vector Regression	3.084	0.454	0.452	0.50 (0.24)
Decision Tree	1.771	0.819	0.819	0.80 (0.33)
Random Forest	1.231	0.913	0.913	0.88 (0.10)
LassoCV	1.065	0.935	0.935	0.90 (0.10)
RidgeCV	1.067	0.935	0.934	0.90 (0.09)

3.3 Grouping students

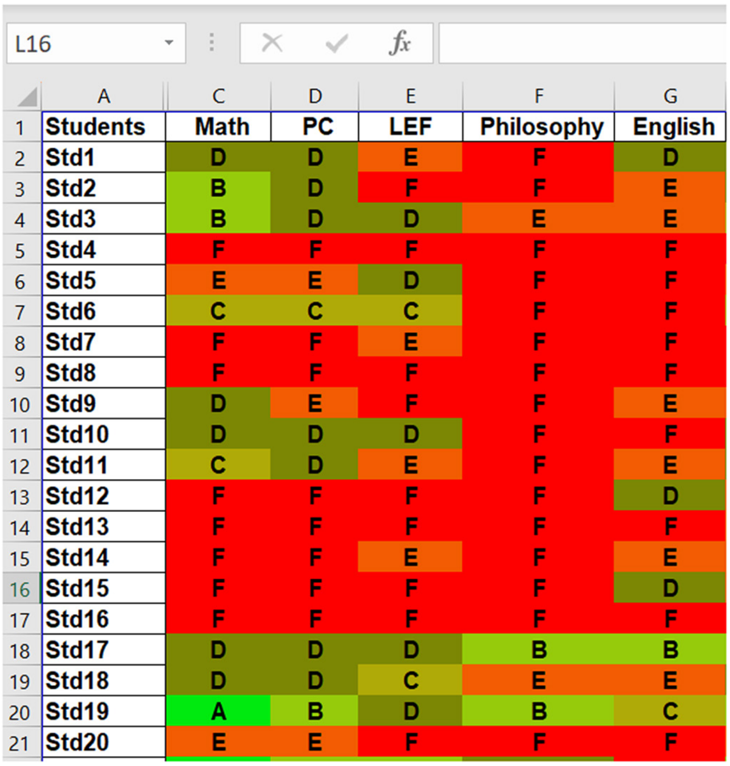
After building the model, the next step is to group students based on their grades. According to the school's pedagogical community, we decided to group students into five (05) different groups, A to F. Table 3 shows the proposed groups and the interval of student grades.

Figure 4 represents an overview of student classification in subjects according to their predicted academic performance. As seen in this figure, 34% of students in Math are arranged in group A, 18% in group B, 7% in group C, 14% in group D, 8% in group E, and 20% in group D. Table 4 presents a part of the student clustering in Math.

Table 3. Group label according to the grade's interval

Groupe	Grade
A	[15 – 20]
B	[12 – 15]
C	[11 – 12]
D	[09 – 11]
E	[07 – 09]
F	<07

Figure 5 represents an overview of student clustering based on their predicted performance in National Exam of Math, PC, LEF, Philosophy, and English. For example, Std1 is classified as D in Math and PC, E in LEF, F in Philosophy, and D in English.



	A	C	D	E	F	G
1	Students	Math	PC	LEF	Philosophy	English
2	Std1	D	D	E	F	D
3	Std2	B	D	F	F	E
4	Std3	B	D	D	E	E
5	Std4	F	F	F	F	F
6	Std5	E	E	D	F	F
7	Std6	C	C	C	F	F
8	Std7	F	F	E	F	F
9	Std8	F	F	F	F	F
10	Std9	D	E	F	F	E
11	Std10	D	D	D	F	F
12	Std11	C	D	E	F	E
13	Std12	F	F	F	F	D
14	Std13	F	F	F	F	F
15	Std14	F	F	E	F	E
16	Std15	F	F	F	F	D
17	Std16	F	F	F	F	F
18	Std17	D	D	D	B	B
19	Std18	D	D	C	E	E
20	Std19	A	B	D	B	C
21	Std20	E	E	F	F	F

Fig. 5. An overview of students grouping by subjects

Table 4. Students grouping in maths

A	B	C	D	E	F
Std19	Std2	Std6	Std1	Std5	Std4
Std21	Std3	Std11	Std9	Std20	Std7
Std26	Std22	Std24	Std10	Std50	Std8
Std27	Std28	Std31	Std17	Std57	Std12
Std29	Std33	Std35	Std18	Std80	Std13
Std40	Std38	Std47	Std23	Std94	Std14
Std44	Std42	Std55	Std25	Std110	Std15
Std72	Std61	Std99	Std48	Std139	Std36
Std77	Std63	Std103	Std51	Std156	Std37
Std84	Std64	Std120	Std52	Std171	Std39
Std87	Std65	Std162	Std56	Std172	Std41
Std96	Std71	Std168	Std66	Std173	Std43
Std97	Std75	Std179	Std78	Std183	Std54
Std98	Std88		Std82	Std201	Std58
Std101	Std93		Std95	Std202	Std68
Std104	Std102		Std105	Std235	Std69
Std106	Std111		Std118		Std74
Std107	Std115		Std119		Std76

4 RECOMMENDATION STRATEGY

In this section, we present a hybrid strategy for generating cooperative learning groups based on learners' predicted performance. In this strategy, we combine student-selected and homogeneous or heterogeneous grouping, depending on various criteria such as the educational objective, goal of creating groups, size, etc. We aim to reap the benefits of letting students choose their collaborators, which enables easier communication, increased comfort in their groups, and heightened enthusiasm for working together. Simultaneously, we aim to leverage the advantages of homogeneity or heterogeneity based on a systematic allocation, such as the predicted performance of students in our case.

4.1 Homogeneity student-selected CLG

The first proposed grouping strategy aims to have students choose the members of their group, not from all class students, but from a homogeneous grouping based on students' academic levels generated from student performance predictions.

For example, creating groups from the E grouping members, as shown in Figure 6. This strategy reduces the competition generated in groups composed of students of different levels. Additionally, instead of treating the classroom as one unit, this strategy helps teachers adapt content according to the group's level, present appropriate pedagogical treatments suitable for these groups (e.g., addressing learning gaps, dealing with advanced concepts, etc.), and consider differences in abilities and interests among their students.

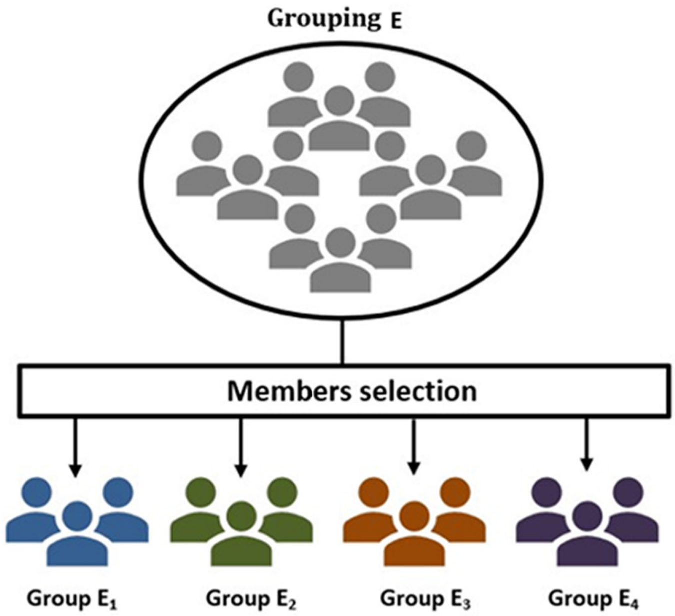


Fig. 6. Homogeneous student-selected CLG

4.2 Heterogeneous student-selected CLG

The second proposed strategy aims to create heterogeneous groups with diverse performance levels. It involves letting students select their group members from the heterogeneous groupings generated based on their performance predictions. For example, creating student-selected groups with one student from the A grouping, two students from E, and two from F (see Figure 7). Students in the A grouping (very efficient grouping) are assigned a tutor role in their group membership. In this case, the teacher assigns a common pedagogical problem to be solved for a determined period. The tutor of each group shares their knowledge and experiences concerning this problem and provides corrective feedback, additional advice, or instruction as needed. This approach promotes the progression of each student within the group, the quality of relationships between learners, their motivation, and the quality of academic learning [4]. For tutors (students in the A grouping in the example), this approach allows them to consolidate their learning by teaching and sharing with their peers, thus building socio-motivational relationships and improving their communication and leadership skills.

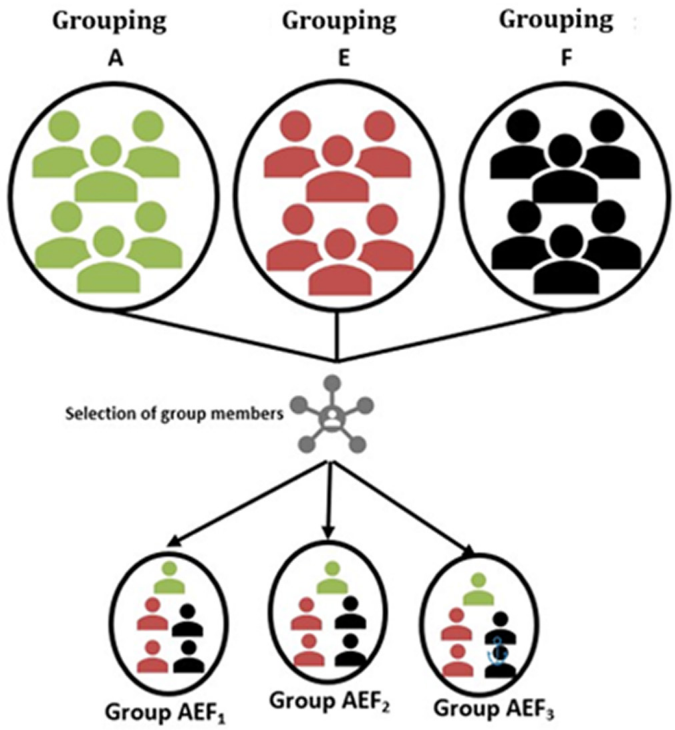


Fig. 7. Heterogeneous student-selected CLG

5 CONCLUSIONS AND FUTURE WORK

In this article, we proposed an approach for recommending cooperative learning groups based on machine learning algorithms. Our study aimed to integrate the most common grouping strategies (student-selected, homogeneous, heterogeneous) with the potential of machine learning, in order to recommend the most optimal group formations based on various criteria, such as educational objectives, group creation goals, and group size. In future work, we plan to investigate the effects of our group recommendations on student achievement and attitudes towards specific subjects, including mathematics, physics, sciences, philosophy, and English, in a selected high school in Morocco.

6 REFERENCES

[1] Hilton, S., & Phillips, F. (2010). Instructor-assigned and student-selected groups: A view from inside. Issues in Accounting Education, 25(1), 15–33. https://doi.org/10.2308/iace.2010.25.1.15

[2] Baum, A., Scarpa, J., Bruzelius, E., Tamler, R., Basu, S., & Faghmous, J. (2017). Targeting weight loss interventions to reduce cardiovascular complications of type 2 diabetes: A machine learning-based post-hoc analysis of heterogeneous treatment effects in the Look AHEAD trial. The Lancet Diabetes & Endocrinology, 5(10), 808–815. https://doi.org/10.1016/S2213-8587(17)30176-6

[3] Bhat, S., Bhat, S., Raju, R., D'Souza, R., & Binu, K. G. (2020). Collaborative learning for outcome based engineering education: A lean thinking approach. Procedia Computer Science, 172, 927–936. https://doi.org/10.1016/j.procs.2020.05.134

[4] Buchs, C. (2008). La distribution des informations dans les dispositifs d'apprentissage entre pairs. Vers des apprentissages en coopération: rencontres et perspectives, 139, 57.

[5] Chapman, K. J., Meuter, M., Toy, D., & Wright, L. (2006). Can't we pick our own groups? The influence of group selection method on group dynamics and outcomes. Journal of Management Education, 30(4), 557–569. https://doi.org/10.1177/1052562905284872

[6] Chen, C. M., & Kuo, C. H. (2019). An optimized group formation scheme to promote collaborative problem-based learning. Computers & Education, 133, 94–115. https://doi.org/10.1016/j.compedu.2019.01.011

[7] 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. https://doi.org/10.1109/ICTAI.2007.155

[8] Cruz, W. M., & Isotani, S. (2014). Group formation algorithms in collaborative learning contexts: A systematic mapping of the literature. In Collaboration and Technology: 20th International Conference, CRIWG 2014, Santiago, Chile, September 7–10, 2014. Proceedings 20 (pp. 199–214). Springer International Publishing. https://doi.org/10.1007/978-3-319-10166-8_18

[9] Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. Three worlds of CSCL. Can we support CSCL, 61491.

[10] Feichtner, S. B., & Davis, E. A. (1984). Why some groups fail: A survey of students' experiences with learning groups. Organizational Behavior Teaching Review, 9(4), 58–73. https://doi.org/10.1177/105256298400900409

[11] Buhmann, A., Paßmann, J., & Fieseler, C. (2020). Managing algorithmic accountability: Balancing reputational concerns, engagement strategies, and the potential of rational discourse. Journal of Business Ethics, 163(2), 265–280. https://doi.org/10.1007/s10551-019-04226-4

[12] Hastie, C. R. (2018). 'TeamUP': An approach to developing teamwork skills in undergraduate midwifery students. Midwifery, 58, 93–95. https://doi.org/10.1016/j.midw.2017.12.026

[13] Herrera-Pavo, M. Á. (2021). Collaborative learning for virtual higher education. Learning, Culture and Social Interaction, 28, 100437. https://doi.org/10.1016/j.lcsi.2020.100437

[14] Marshall, L. L., & Bolt-Lee, C. E. (2022). A neuroeducation and cognitive learning-based strategy to develop active learning pedagogies for accounting education. Account. Educ. J., 32, 2022.

[15] Johnson, D. W., & Johnson, R. T. (1987). Learning together and alone: Cooperative, competitive, and individualistic learning. Prentice-Hall, Inc.

[16] Kagan, S., & Kagan, M. Kagan Cooperative Learning (2009).

[17] Kooloos, J. G., Klaassen, T., Vereijken, M., Van Kuppeveld, S., Bolhuis, S., & Vorstenbosch, M. (2011). Collaborative group work: Effects of group size and assignment structure on learning gain, student satisfaction and perceived participation. Medical Teacher, 33(12), 983–988. https://doi.org/10.3109/0142159X.2011.588733

[18] Krouska, A., & Virvou, M. (2019). An enhanced genetic algorithm for heterogeneous group formation based on multi-characteristics in social-networking-based learning. IEEE Transactions on Learning Technologies, 13(3), 465–476. https://doi.org/10.1109/TLT.2019.2927914

[19] Lejk, M., Wyvill, M., & Farrow, S. (1999). Group assessment in systems analysis and design: A comparison of the performance of streamed and mixed-ability groups. Assessment & Evaluation in Higher Education, 24(1), 5–14. https://doi.org/10.1080/0260293990240101

[20] Maina, E. M., Oboko, R. O., & Waiganjo, P. W. (2017). Using machine learning techniques to support group formation in an online collaborative learning environment. International Journal of Intelligent Systems & Applications, 9(3), 26–33. https://doi.org/10.5815/ijisa.2017.03.04

[21] Moreno, J., Ovalle, D. A., & Vicari, R. M. (2012). A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. Computers & Education, 58(1), 560–569. https://doi.org/10.1016/j.compedu.2011.09.011

[22] Nhan, H., & Nhan, T. A. (2019). Different grouping strategies for cooperative learning in English majored seniors and juniors at can tho university, Vietnam. Education Sciences, 9(1), 59. https://doi.org/10.3390/educsci9010059

[23] Pang, Y., Xiao, F., Wang, H., & Xue, X. (2014, December). A clustering-based grouping model for enhancing collaborative learning. In 2014 13th International Conference on Machine Learning and Applications (pp. 562–567). IEEE. https://doi.org/10.1109/ICMLA.2014.94

[24] Pasina, I., Bayram, G., Labib, W., Abdelhadi, A., & Nurunnabi, M. (2019). Clustering students into groups according to their learning style. MethodsX, 6, 2189–2197. https://doi.org/10.1016/j.mex.2019.09.026

[25] Roger, T., & Johnson, D. W. (1994). An overview of cooperative learning. Creativity and Collaborative Learning, 1–21.

[26] Sanz-Martínez, L., Martínez-Monés, A., Bote-Lorenzo, M. L., Munoz-Cristóbal, J. A., & Dimitriadis, Y. (2017). Automatic group formation in a MOOC based on students' activity criteria. In Data driven approaches in digital education: 12th European conference on technology enhanced learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017, Proceedings 12 (pp. 179–193). Springer International Publishing. https://doi.org/10.1007/978-3-319-66610-5_14

[27] Theobald, E. J., Eddy, S. L., Grunspan, D. Z., Wiggins, B. L., & Crowe, A. J. (2017). Student perception of group dynamics predicts individual performance: Comfort and equity matter. PloS one, 12(7), e0181336. https://doi.org/10.1371/journal.pone.0181336

[28] van der Laan Smith, J., & Spindle, R. M. (2007). The impact of group formation in a cooperative learning environment. Journal of Accounting Education, 25(4), 153–167. https://doi.org/10.1016/j.jaccedu.2007.09.002

[29] Qazdar, A., Er-Raha, B., Cherkaoui, C., & Mammass, D. (2019). A machine learning algorithm framework for predicting students performance: A case study of baccalaureate students in Morocco. Educ. Inf. Technol., 24(6), 3577–3589. https://doi.org/10.1007/s10639-019-09946-8

[30] El Guabassi, I., Bousalem, Z., Marah, R., & Qazdar, A. (2021). A recommender system for predicting students' admission to a graduate program using machine learning algorithms. International Journal of Online and Biomedical Engineering (iJOE), 17(02), 135–147. https://doi.org/10.3991/ijoe.v17i02.20049

[31] El Guabassi, I., Bousalem, Z., Marah, R., & Qazdar, A. (2021). Comparative analysis of supervised machine learning algorithms to build a predictive model for evaluating students' performance. International Journal of Online and Biomedical Engineering (iJOE), 17(02), 90–105. https://doi.org/10.3991/ijoe.v17i02.20025

[32] Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. Computers in Human Behavior, 104, 106189. https://doi.org/10.1016/j.chb.2019.106189

[33] El Guabassi, I., Al Achhab, M., Jellouli, I., & Mohajir, B. E. E. (2018). Personalized ubiquitous learning via an adaptive engine. International Journal of Emerging Technologies in Learning (iJET), 13(12), 177–190. https://doi.org/10.3991/ijet.v13i12.7918

7 AUTHORS

Zakaria Bousalem holds a Ph.D in Computer Science from in the Faculty of Sciences and Technics, University Hassan 1st, Settat, Morocco. He has been working for a series of years as a computer science teacher in public sector secondary education. His research interests are artificial intelligence in education (AIEd), Big Data, NoSQL databases, data integration, XML processing, relational database management systems (RDBMS), query processing and optimization.

Aimad Qazdar Assistant Professor at the National School of Applied Sciences (ENSA) - Agadir, Ibn Zohr University. From 2019 to 2022, he was an Assistant Professor at the Faculty of Sciences Semlalia – Cadi Ayyad University in Marrakech Morocco. Ex-computer science teacher in high school for 8 years. He holds a Ph.D. degree in Computer Science in 2017, as well as a Master's degree in Software Engineering and computer networks in 2010 from the Ibn Zohr University, National School of Applied Sciences (ENSA) Agadir, Morocco. His research focuses on the study and use of emerging digital technologies (Big Data, IoT, IA, Cloud, VR, AR, etc.) for the development and deployment of value-added solutions in digital transformation of two fields of application: Education (University 4.0) and the industrial field (Industry 4.0).

Inssaf El Guabassi holds a Ph.D from the Faculty of Sciences, Abdelmalek Essaadi University, Tetouan, Morocco and an engineering degree in computer science from the National School of Applied Sciences (ENSA) Tetuan, Morocco. She is currently an Assistant Professor in the department of computer science at the Faculty of Sciences and affiliated with the LAROSERI Laboratory at Chouaib Doukkali University, El Jadida, Morocco. Her research interests are artificial intelligence in education (AIEd), adaptive learning systems and ubiquitous learning.