



HAL
open science

Aide au suivi de la progression de groupes d'apprenants pour la mise en place d'une pédagogie différenciée

Fatima Harrak, François Bouchet

► To cite this version:

Fatima Harrak, François Bouchet. Aide au suivi de la progression de groupes d'apprenants pour la mise en place d'une pédagogie différenciée. 10e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain, Marie Lefevre, Christine Michel, Jun 2021, Fribourg, Suisse. pp.312-317. hal-03287778

HAL Id: hal-03287778

<https://hal.science/hal-03287778v1>

Submitted on 15 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Aide au suivi de la progression de groupes d'apprenants pour la mise en place d'une pédagogie différenciée

Fatima Harrak et François Bouchet

Sorbonne Université, CNRS, LIP6, F-75005 Paris, France
fatima.harrak@lip6.fr
francois.bouchet@lip6.fr

Résumé. Ce travail se situe dans le cadre d'un assistant pédagogique à destination des professeurs des écoles pour accompagner leurs élèves de primaire en mathématiques. Pour proposer à l'enseignant une vision synthétique des activités de sa classe, nous utilisons des méthodes de clustering pour constituer automatiquement des groupes d'élèves de sa classe en fonction des activités réalisées dans l'application, où chaque élève suit un parcours individuel. Après une présentation des traces utilisées et de la qualité intrinsèque des clusters produits (coefficient de silhouette), nous analysons les retours d'une enquête auprès de 32 enseignants utilisateurs pendant au moins 6 semaines. Ces analyses ouvrent des perspectives quant aux améliorations possibles des clusters et de leur visualisation.

Mots-clé: Clustering · traces d'activités · tableau de bord.

Abstract. This work deals with a pedagogical assistant for primary school teachers to assist students in mathematics. To give the teacher a synthetic vision, we use clustering methods to automatically provide groups of students in their class based on their activities in the application, where each student follows a different path. We first present the traces used and the intrinsic quality of the clusters (Silhouette coefficient), and then the answers to a questionnaire filled by 32 teachers who have used the application for at least 6 weeks. These analyses open perspectives on how to improve the clusters and their visualization.

Keywords: Clustering · Students' activities · Dashboard

1 Introduction

La pédagogie différenciée est “une pédagogie fondée sur la différenciation des processus d'apprentissage” [8] dont la mise en oeuvre pour l'enseignant peut être complexe car elle suppose de différencier les contenus et de proposer des structures de travail dynamiques adaptées aux savoirs et lacunes de chacun [7].

Grâce aux algorithmes d'apprentissage adaptatif, de plus en plus d'applications proposent un parcours personnalisé à chaque élève. Cependant cela complique

pour l'enseignant le suivi des parcours individuels. Il faut donc lui proposer une vision synthétique (via un tableau de bord) du travail effectué par ses élèves, mais aussi lui permettre de déterminer des groupes d'élèves avec des lacunes communes auxquels proposer une activité de remédiation commune. Notre objectif est donc de répondre aux questions de recherche suivantes : (QR1) Comment regrouper des élèves ayant des parcours d'apprentissage hétérogènes dans des dispositifs d'apprentissage ? et (QR2) Comment restituer ces groupes de manière compréhensible et acceptable par les enseignants ?

2 Travaux connexes

Des recherches antérieures ont largement proposé des restitutions aux enseignants à l'échelle de groupes via des méthodes de clustering pour aider au suivi de progression des élèves. Ainsi [3] propose un système adaptatif pour constituer des groupes d'élèves de collège de différents niveaux de compétences en lecture. Les résultats du clustering (K-Means) ont montré des corrélations positives entre les niveaux de difficulté du texte et le temps de lecture pour les 4 groupes étudiés. Li et Yoo ont proposé une chaîne de Markov basée sur le clustering pour modéliser les comportements en ligne des élèves pendant le processus d'apprentissage pour un enseignement plus efficace et adaptatif [5].

D'autres approches utilisent le clustering pour regrouper les séquences d'interactions des élèves dans les environnements hybrides [1]. McBroom et al. [6] s'appuient sur un algorithme de clustering hiérarchique pour identifier les tendances comportementales d'élèves à partir de données temporelles, notamment en termes d'abandon par types d'exercices, information qui pourrait nous intéresser pour analyser les trajectoires individuelles.

Par ailleurs, K-Means a été utilisé dans un contexte différent pour faire des regroupements d'élèves selon la nature des questions posées [4]. Dans la suite de cet article, nous utilisons également K-Means pour faire des regroupements des élèves ayant des parcours d'apprentissage très hétérogènes.

3 Contexte

3.1 L'application Adaptiv'Math

Adaptiv'Math est un assistant pédagogique numérique en mathématiques pour les enseignants et élèves du cycle 2 (CP à CE2). Il associe deux algorithmes d'intelligence artificielle : (1) pour des parcours individualisés (ZPDES [2]) et (2) pour synthétiser les expériences individuelles de chaque élève. L'enseignant contrôle ces algorithmes via son tableau de bord. L'application permet aux élèves de travailler sur 5 modules de mathématiques. Chaque module est découpé en objectifs (15 à 20 par module) visant à développer une compétence particulière par des exercices de difficulté progressive, eux-mêmes découpés en 4 à 6 niveaux. Le fonctionnement de l'application au sein d'un module se résume en 5 étapes :
 (1) Un test diagnostique (40 questions) évalue le niveau initial de l'élève.
 (2) Un premier regroupement d'élèves par clustering est proposé à l'enseignant.
 (3) Chaque élève réalise une séquence d'exercices (cf. Figure 1) où chaque exercice

est choisit au fur et à mesure en fonction des compétences jugées acquises.

(4) Les algorithmes de clustering analysent les résultats de tous les élèves pour proposer de nouveaux regroupements.

(5) L'analyse des résultats, dont les clusters, est transmise aux enseignants via un tableau de bord. Les élèves ont également des retours d'information.

Encore en phase de déploiement, l'application a été utilisée d'octobre à décembre 2020 par 137 classes. 119 classes ont effectué le test initial sur au moins un module (étape 2 atteinte) et 102 ont travaillé sur l'application au-delà de ce test (étape 3 atteinte). Deux modules sur 5 étaient présents et nous nous concentrons dans la suite sur le module 1 (ayant donné lieu à plus d'interactions).

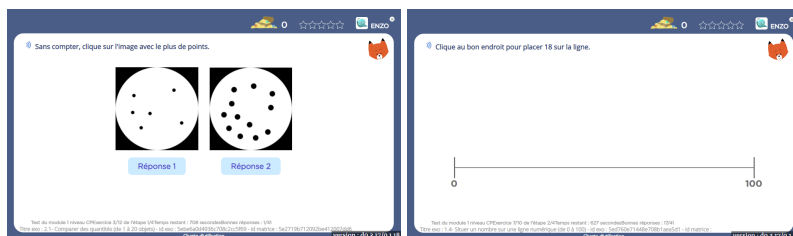


Fig. 1. Deux exemples d'exercices du module 1 sur le sens des nombres demandant de choisir le nuage avec le plus de points et de placer 18 sur une ligne allant de 0 à 100

3.2 Données collectées

Les traces d'activités de l'application sont étroitement liées aux "gameplays", c'est-à-dire un patron d'exercices (ex : QCM, ligne numérique ou glisser-déposer) portant sur différents objectifs. Au sein d'un gameplay, différents *événements* peuvent avoir lieu correspondant à des moments-clé de l'application (ex : un exercice est chargé) ou à une action de l'apprenant (ex : une réponse est sélectionnée). Chaque événement dispose d'un ensemble d'*attributs* associés. Par exemple, lorsqu'un élève termine un exercice, un événement `EvtExerciseEnded` est généré avec comme attributs associés la date et heure (timestamp), l'ID d'exercice, l'ID d'élève, le score sur l'exercice (normalisé entre 0 et 1) et le temps en millisecondes depuis l'événement `EvtExerciseStarted` associé.

Au sein du module 1, nous avons collecté 1 048 575 traces d'activités correspondant à 1 583 exercices différents réalisés par 1 534 élèves. Toutes les données remontées sont complètement anonymisées.

4 Constitution de groupes

Principes de l'algorithme. Bien qu'il soit possible d'appliquer un clustering aux données d'une classe, le nombre réduit de traces par élève limite la qualité des clusters obtenus. Nous proposons donc une approche en deux temps :

(1) phase d'*apprentissage* : on entraîne l'algorithme pour déterminer le meilleur partitionnement à partir des traces de toutes les classes à un instant T_{app}

(2) phase de *prédiction* : le partitionnement précédemment appris est appliqué pour prédire l'appartenance d'un nouvel élève à un cluster à un instant $T_{pred} > T_{app}$, selon la distance de celui-ci avec chacun des centroïdes des clusters.

On utilise ainsi indirectement le clustering à des fins de prédictions, tel que fait dans [4]. La nature des exercices étant différente dans chaque module, il y a un clustering séparé pour chacun des modules de l'application.

Le premier clustering sur un module est donc réalisé exclusivement sur les données du test initial (cf. étape 2) avec les mêmes exercices travaillés par tous les élèves. Par la suite (cf. étape 4), les clusters produits dépendent de l'activité des élèves, et indirectement de la nature de ces exercices. Nous avons également intégré des heuristiques de sélection automatique du nombre de clusters et garantissant un nombre de clusters "raisonnable" pour l'enseignant (entre 2 et 5).

Prétraitement et application. Tout d'abord, il faut passer de la représentation événementielle des traces à une représentation matricielle où chaque ligne représente un élève et chaque colonne représente la valeur d'une variable donnée. Ensuite, nous avons calculé pour chaque objectif : le score moyen obtenu, la durée passée sur un exercice, le niveau de difficulté maximum atteint et le nombre d'exercices travaillés. Puis un ensemble de prétraitements est effectué : remplacement des données aberrantes, standardisation et réduction d'attributs par ACP. Enfin nous avons entraîné un algorithme K-Means++ (1000 fois pour éviter les optimums locaux) sur l'ensemble des données du module jusqu'à une date fixe en considérant les attributs ci-dessus, puis choisit le partitionnement par la méthode du coude (elbow method) en ayant fait varier le nombre de clusters de 2 à 20.

Résultats. Pour le module 1, nous avons trouvé 6 clusters globaux de référence à partir des traces d'activités de 1534 élèves (cf. Table 1). Nous avons utilisé le coefficient de silhouette (variant entre -1 et 1) pour évaluer la qualité des clusters. Quand le coefficient est proche de 0 mais positif, les clusters "se touchent" et des individus sont proches des centroïdes de deux clusters différents mais sans chevauchement. Les coefficients de silhouette des 6 clusters de référence variaient entre -0.027 et 0.21 (0.093 en moyenne). Ce score proche de 0 s'explique par le grand nombre d'attributs utilisés (40), même après l'ACP, qui rend plus difficile la séparation des clusters. On regroupe en effet des élèves effectuant des exercices différents ; sur le test initial uniquement (étape 2 - exercices identiques pour tous) les coefficients de silhouette sont logiquement supérieurs (0.23 en moyenne).

Nous avons considéré 5 variables descriptives pour caractériser les clusters (cf. Table 1) et 3 groupes distincts apparaissent : un groupe (1, 3 et 6) avec des scores et nombres d'exercices/objectifs plus élevé et moins rapides que les autres. Un autre groupe (2 et 4) à l'extrême des caractéristiques du premier groupe et un groupe (5) plus rapide mais entre les deux autres en termes de score, nombre d'exercices, etc. Le fait que 3 groupes ressortent et non 6 ne veut pas dire qu'un regroupement en 3 clusters serait plus pertinent, mais que les variables explicatives choisies ne permettent pas de faire ressortir ce qui distingue les clusters 2 et 4 par exemple. Ces variables doivent en effet être suffisamment

compréhensibles par les enseignants ce qui demande un équilibre. Dans le tableau de bord testé, seules 3 de ces 5 variables étaient présentées (cf. Figure 2).

Tableau 1. Moyennes des 6 clusters globaux du module 1 selon 5 variables descriptives

Cluster	N	Score	Durée	Nb objectifs	Nb exerc. faits	Nb exerc. réussis
cluster1	124	0.65	758.73	10.82	105.59	65.46
cluster2	503	0.55	153.04	4.64	41.56	22.47
cluster3	36	0.69	862.05	11.92	210.58	144.47
cluster4	504	0.58	127.84	5.67	49.49	28.02
cluster5	111	0.61	37.22	8.14	74.64	43.05
cluster6	256	0.66	848.76	9.05	103.11	66.11

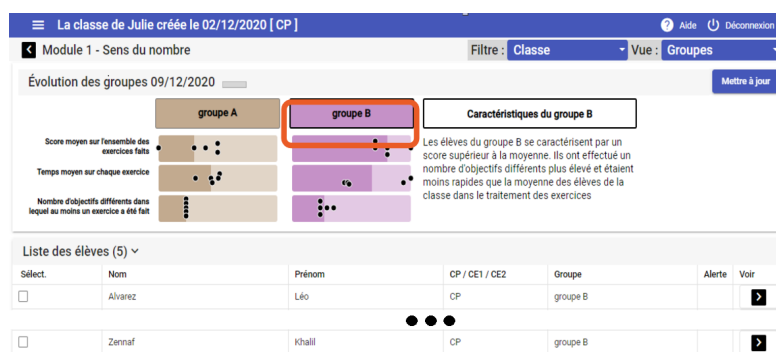


Fig. 2. Tableau de bord enseignant visualisant les clusters. Jauge = niveau moyen de chaque groupe d'élèves (ici A et B), point = un élève. Un clic sur un point le sélectionne dans la liste en-dessous, un clic sur un groupe affiche ses caractéristiques (ici le B).

5 Retour des enseignants

Pour répondre à la QR2, un questionnaire a été envoyé début décembre 2020 auquel 29 enseignants (sur 36) ont répondu afin d'évaluer l'accessibilité de la visualisation (cf. Figure 2), leur bonne compréhension de ce qui est montré et l'intérêt de cette information pour mettre en place une pédagogie différenciée.

Accessibilité. 18 enseignants déclarent être allés voir la visualisation de groupe générée par le clustering, 9 y sont parvenus facilement sans aide, 7 facilement avec les fiches fournies et 2 avec difficulté malgré l'aide. Il y a donc une difficulté d'accès à cette interface, qui a limité la visibilité des résultats proposés, d'autant que 7 des 11 non répondants déclarent avoir tenté d'y accéder sans succès.

Compréhensibilité de la visualisation des clusters. Parmi les 18 enseignants qui y ont eu accès, 5 l'ont comprise très facilement, 9 facilement avec les fiches fournies, 2 difficilement avec cette même aide, et 2 n'y sont pas parvenus faute de bien comprendre les critères retenus, soulignant le besoin d'explicitier davantage ces choix. Une majorité semble comprendre globalement la démarche et l'intérêt. Ce point ressort aussi quand les enseignants notent (de 0 à 10) l'utilité des critères de description fournis ($M = 6$, $\sigma = 2$) par rapport à leurs besoins.

Utilité pour une pédagogie différenciée. 25 enseignants répondent "oui" (1 "non", 3 "je ne sais pas") montrant qu'en dépit des problèmes d'accès et de lisibilité pour certains, la majorité perçoit l'intérêt du clustering et des groupes.

6 Conclusion

Nous avons présenté ici un modèle de clustering entraîné sur des sessions antérieures d'un module d'exercices de mathématiques, afin de prédire le cluster d'appartenance de nouveaux élèves à partir de leurs traces d'activités. Nous avons montré que les clusters obtenus, bien que proches dans l'espace, sont distincts et globalement explicables de manière compréhensible pour des enseignants, bien qu'il soit difficile de caractériser certains avec les variables simples utilisées. Ce choix de variables est un des points faibles actuellement mis en avant dans les retours d'enseignants pour la compréhension des groupes, et constitue donc un des points majeurs d'amélioration en collaboration avec eux. Une autre limite de notre approche est liée au démarrage à froid : il faut disposer d'un volume de données important et représentatif de l'ensemble des élèves pour constituer les clusters globaux de référence. Ces clusters peuvent ensuite être réappris au fil du temps quand les données augmentent. Actuellement, ce problème est contourné en proposant un clustering par classe (sur des données locales), non optimal, pendant les premières semaines d'ouverture d'un module.

7 Remerciements

Nous remercions EvidenceB pour la diffusion du questionnaire et le Partenariat d'Innovation sur l'IA (P2IA) de la DNE pour le financement de ce projet.

Références

1. Akpınar, N.J., Ramdas, A., Acar, U.: Analyzing student strategies in blended courses using clickstream data. arXiv preprint arXiv:2006.00421 (2020)
2. Clément, B., Roy, D., Oudeyer, P.Y., Lopes, M.: Online Optimization of Teaching Sequences with Multi-Armed Bandits (2014), <https://hal.inria.fr/hal-01016428>
3. Fadljević, L., Maitz, K., Kowald, D., Pammer-Schindler, V., Gasteiger-Klicpera, B.: Slow is good: the effect of diligence on student performance in the case of an adaptive learning system for health literacy. In: Proc. Tenth Int. Conf. on Learning Analytics & Knowledge. pp. 112–117. LAK '20, ACM, New York, NY, USA (Mar 2020). <https://doi.org/10.1145/3375462.3375502>
4. Harrak, F., Bouchet, F., Luengo, V.: From Students' Questions to Students' Profiles in a Blended Learning Environment. *Journal of Learning Analytics* **6**(1), 54–84 (Apr 2019). <https://doi.org/10.18608/jla.2019.61.4>
5. Li, C., Yoo, J.: Modeling student online learning using clustering. In: Proceedings of the 44th annual Southeast regional conference. pp. 186–191 (2006)
6. McBroom, J., Yacef, K., Koprinska, I.: DETECT: A Hierarchical Clustering Algorithm for Behavioural Trends in Temporal Educational Data. In: Artificial Intelligence in Education. pp. 374–385. LNCS, Springer (2020). https://doi.org/10.1007/978-3-030-52237-7_30
7. Meirieu, P.: L'École, mode d'emploi: Des méthodes actives à la pédagogie différenciée. ESF éditeur, 15^{ème} edn. (2009)
8. Przesmycki, H.: Pédagogie différenciée. Hachette, Paris (1991)