

APPORT DES *LEARNING ANALYTICS*

Geoffroy BONNIN & Anne BOYER

Les *Learning Analytics* ou analyse de l'apprentissage constituent une discipline émergente à la confluence de l'informatique, des sciences de l'éducation et des mathématiques. Leur objet d'étude est la collecte, l'analyse et l'utilisation intelligentes de données produites par l'apprenant. Si les *Learning Analytics* puisent leurs techniques dans plusieurs communautés, elles constituent en tant que telles un phénomène assez récent apparu avec la généralisation du numérique éducatif et la disponibilité de données massives sur l'apprentissage. Cet article présente brièvement leurs objectifs et quelques exemples illustratifs de leurs apports potentiels à l'amélioration ou à la meilleure compréhension de l'apprentissage.

LEARNING ANALYTICS : DEFINITION ET OBJECTIFS

La multiplication des contenus pédagogiques en ligne, des plateformes d'apprentissage et des services numériques, la généralisation de l'usage du numérique dans l'éducation et l'apparition de technologies permettant la capture, le stockage, la recherche, le partage, l'analyse et la visualisation de données massives (*Big Data*) ont permis la constitution de corpus massifs de données sur les comportements des apprenants. Ce contexte a permis l'émergence des *Learning Analytics* – ou analyse de l'apprentissage. Les *Learning Analytics* sont une discipline récente qui s'intéresse à la collecte, l'analyse, la modélisation, la synthèse et la communication des données relatives aux apprenants et à leurs environnements, afin de mieux comprendre et améliorer l'apprentissage. Les *Learning Analytics* collectent les traces numériques laissées par les apprenants, les exploitent pour améliorer l'apprentissage et permettent ainsi une approche des sciences de l'apprentissage dirigées par les données qui s'applique autant aux situations de formation formelles qu'informelles. Les notions de protection de la vie privée (*privacy*) et d'éthique sont au centre des approches développées et des dispositifs mis en œuvre. Si les *Learning Analytics* utilisent des techniques issues de plusieurs communautés comme la fouille de données, la visualisation de l'information, le traitement automatique du langage, la recommandation sociale, etc., elles sont désormais un domaine de recherche à part entière qui a notamment sa propre conférence internationale – LAK pour *Learning Analytics and Knowledge* – dont la première édition s'est déroulée en 2011 à Banff au Canada.

Les applications possibles des *Learning Analytics* sont nombreuses et variées puisqu'elles vont de la prédiction de l'abandon dans les MOOC à la personnalisation des apprentissages, de l'amélioration de la qualité de l'apprentissage à la recommandation personnalisée de ressources de formation, de l'aide à la construction d'un parcours d'apprentissage à la détection des décrocheurs, etc. L'exploitation des traces d'interaction en situation d'apprentissage permet différents types de retours d'expérience, vers l'apprenant, le tuteur ou le formateur, les pairs, le concepteur de la ressource ou l'institution, le plus souvent sous la forme synthétique d'indicateurs ou de recommandations personnalisées. Les *Learning Analytics* ne se contentent pas d'analyser les résultats de l'apprentissage, mais fournissent une vision globale et dynamique du processus d'apprentissage ainsi que des informations pertinentes pour l'aide à la décision dans l'établissement des politiques éducatives.

QUELQUES EXEMPLES

Recommandation personnalisée de ressources pédagogiques

La mise en place du LMD a introduit des dimensions nouvelles dans les formations comme l'individualisation des parcours de formation. Selon (Sauve et al., 1993), permettre l'individualisation d'un parcours consiste à mettre en place un environnement éducatif qui tient compte des caractéristiques individuelles de l'apprenant comme par exemple son niveau, ses habitudes, ses besoins, ses objectifs, sa langue, sa culture, etc. Parmi les composantes de cet environnement éducatif figurent les contenus. Il s'agit de faire en sorte

que, une fois les objectifs fixés, chaque apprenant puisse disposer des contenus qui lui conviennent le mieux.

Le projet du programme d'investissement d'avenir (PIA 1 e-éducation) Périclès conçoit entre autres un outil permettant de suggérer à chaque apprenant des compléments de cours ou des activités pédagogiques, en fonction de ses parcours, niveau et objectif. L'objectif est de recommander des ressources adaptées à chaque apprenant en s'appuyant, d'une part sur les évaluations et autres traces de navigation laissées par l'apprenant, et d'autre part sur les choix réalisés par les apprenants de profil similaire. Un deuxième objectif concerne l'identification de parcours type à partir des pratiques effectives afin de les confronter à ceux imaginés par les formateurs. De plus, la détermination automatique de communautés homogènes d'apprenants (communautés de pratiques, d'intérêt...) participe à une meilleure compréhension des populations étudiantes et permet de cartographier le public étudiant d'un établissement en fonction de ses usages observés.

Périclès (Chan et al., 2014) propose deux scénarios, chacun reflétant un usage de la recommandation personnalisée. Le premier scénario (recommandation publique – utilisateur non identifié) se concrétise par une interface disponible sur un portail de consultation de ressources comme celui d'une université numérique thématique. Cette interface se présente comme une composante autonome agrégée dans le portail et est liée à un logiciel de recommandations personnalisées de ressources pédagogiques. Le deuxième scénario (recommandation privée – utilisateur connecté) est destiné à être intégré dans la plateforme numérique d'apprentissage (*Learning Management System*) d'une université ou d'un organisme de formation. Dans ce cas, chaque étudiant peut bénéficier d'un service de recommandation personnalisée, en lien avec son cursus de formation, matérialisé sous forme d'une interface intégrée à son environnement pédagogique privé.

Aménagement et décoration de salles de classe

L'aménagement et la décoration des salles de classe peuvent avoir une grande influence sur l'apprentissage des élèves, en particulier pour les élèves du primaire (Lim et al., 2012). Partant de ce constat, beaucoup de travaux se sont intéressés à la manière de disposer les bureaux, à l'apparence visuelle de la classe, son niveau de coloration, la quantité de dessins accrochés au mur, etc. Jusqu'à présent, la stratégie était d'étudier l'influence de propriétés particulières de salles de classe sur les résultats des élèves. L'émergence des *Learning Analytics* a permis une ouverture vers une nouvelle stratégie : rechercher des tendances dans les choix d'aménagement et de décoration faits par les enseignants.

(Almeda et al., 2014) est un exemple de travail qui suit cette stratégie. Les auteurs ont pris des photos de 30 salles de classes et ont encodé leurs propriétés en fonction de diverses catégories comme planning, compétences, devoirs, règlement, décoration, comportement, etc. Ils utilisent la somme des éléments de chaque catégorie comme coordonnées dans un espace vectoriel dont les dimensions sont les catégories elles-mêmes et appliquent un algorithme de *clustering* standard pour déterminer ce qui distingue les salles de classe les unes des autres. Les auteurs ont ainsi découvert que le contenu des éléments affichés est fortement discriminant, ce qui semble refléter les différences dans les approches pédagogiques des enseignants.

Détection d'élèves à risque

Les systèmes de notification de situation d'échec potentiel d'élèves à destination des enseignants sont un autre exemple d'outils destinés à améliorer l'apprentissage, qui sont de plus en plus populaires au sein des universités.

Un exemple d'un tel système est le *Early Warning System* (EMS) de (Lonn et al., 2012). Ce système a été pensé pour le domaine des STEM (*Science, Technology, Engineering and Mathematics*). Le système utilise des combinaisons simples des notes des étudiants et de fréquence de connexion sur le site du cours pour prédire les étudiants en risque d'échec.

Pour déterminer les meilleurs paramètres et seuils à utiliser pour identifier de tels apprenants, les concepteurs du logiciel ont interagi de manière poussée avec les enseignants. Une première expérimentation de cet outil a permis une amélioration des performances des étudiants (Krumm et al., 2014).

Prédiction d'abandon dans les MOOC

Ces dernières années, les MOOCs (Massive Open Online Courses) sont devenus un dispositif très populaire de l'e-éducation et ont attiré une attention considérable de la part de diverses institutions universitaires à travers le monde. Un MOOC offre non seulement de nouvelles manières d'acquérir des connaissances, mais permet aussi de récupérer des données d'usage considérables qui peuvent être exploitées pour améliorer la qualité de l'enseignement.

Une des différences majeures avec les outils pédagogiques habituels est que les MOOC sont libres d'accès. La conséquence principale est qu'en général plus de 90% des inscrits abandonnent les activités proposées en cours de route (Breslow et al., 2013). Éviter les décrochages est donc un des défis principaux de cet outil.

Une étude de Stanford montre que certaines interventions particulières peuvent aider à réduire le taux de décrochage, à condition de détecter suffisamment tôt les utilisateurs susceptibles d'un tel décrochage (Halawa et al., 2014). Les auteurs proposent un modèle capable de prédire les abandons en fonction des usages des étudiants au sein du MOOC. Leur modèle a obtenu des résultats prometteurs puisqu'il est capable de prédire les abandons dans près de la moitié des cas, alors même que les étudiants sont encore actifs.

Tableaux de bord

Les tableaux de bord ont une grande popularité en *Learning Analytics*. De nombreuses études ont montré que leur utilisation peut grandement améliorer les résultats des apprenants (Arnold et al., 2010 ; Thille et al., 2011).

Ces tableaux de bords peuvent par exemple consister en des cartes de compétences permettant aux élèves de visualiser leur apprentissage, comme proposé par (Gran et al., 2014). Ce tableau de bord est à destination d'étudiants du *Master of Business Administration* (MBA) et représente leur apprentissage relativement aux connaissances et compétences qu'ils sont censés acquérir : capacité à intégrer des concepts fondamentaux du marketing, capacité à communiquer de manière professionnelle, etc. Les données utilisées pour établir les cartes de compétences proviennent des nombreuses évaluations des enseignants. Une première expérimentation impliquant plus de 200 étudiants et 16 cours différents montre que les étudiants ayant pu utiliser ce tableau de bord sont beaucoup plus nombreux à se réinscrire le semestre suivant. Cependant, aucune différence statistiquement significative n'a pu être trouvée pour ce qui est de leur réussite aux examens finaux.

Un autre exemple de tableau de bord est celui du projet eLAT (*exploratory Learning Analytics Toolkit*) de (Dyckhoff, 2014). Ici, le but est d'aider les enseignants à la réflexion et à l'amélioration de leurs méthodes d'enseignement en leur permettant d'explorer et de mettre en relation leurs propres usages, les usages des étudiants et leurs résultats via une interface utilisant des indicateurs graphiques. Le projet est le résultat d'un développement itératif sur plusieurs années. La Figure 1 donne un aperçu de l'un de ses indicateurs graphiques. Elle présente un diagramme à bulles qui met en relation les résultats des étudiants avec des scores d'activité calculés en fonction de traces d'usage pour différents domaines (une couleur par domaine). Le système a été expérimenté sur 6 classes de plus de 100 étudiants. Les enseignants ont pu l'utiliser sur une durée de 8 à 10 semaines. Les résultats montrent que l'utilisation de ce tableau de bord a grandement facilité la tâche des enseignants dans leur réflexion sur leurs cours et sur leurs étudiants.

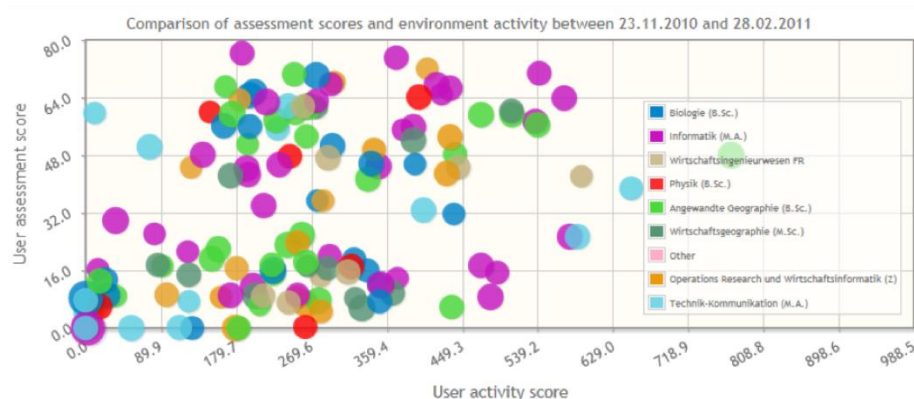


Figure 1 : Exemple d'indicateurs (Dyckhoff, 2014)

CONCLUSION

Les applications des *Learning Analytics* sont multiples et variées, et leur potentiel considérable. Les recherches, essentiellement pluri-disciplinaires, se développent et se diversifient. Si elles s'intéressent toujours à la prédiction de l'abandon des apprenants et à la performance académique, elles travaillent aussi à mieux comprendre comment les étudiants ou les élèves apprennent et par là-même améliorer les pratiques pédagogiques, la conception de ressources éducatives, la personnalisation des apprentissages ou l'animation de communautés.

REFERENCES

- Almeda, V., Scupelli, P., Baker, R., Weber et M., Fisher, A. (2014). Clustering of design decisions in classroom visual displays. *Proc. LAK*, p. 44-48.
- Arnold, K. et Pistilli, M. (2012). Course signals at purdue: Using learning analytics to increase student success. *Proc. LAK*.
- Barrett, P., Zhang, Y., Moffat, J. et Kobbacy, K. (2013). A holistic, multi-level analysis identifying the impact of classroom design on pupils' learning. *Building and Environment*, 59, p. 678-689.
- Nguyen, C., Roussanaly A. et Boyer, A. (2014). Learning Resource Recommendation: An Orchestration of Content-Based Filtering, Word Semantic Similarity and Page Ranking, *Proc. EC-TEL*.
- Dyckhoff, A. (2014). Action Research and Learning Analytics in Higher Education. *Thèse de doctorat de l'Université technique de Rhénanie-Westphalie, Aix-la-Chapelle*.
- Grann, J., et Bushway, D. (2014). Competency Map: Visualizing Student Learning to Promote Student Success. *Proc. LAK*, p. 168-172.
- Halawa, S., Greene, D. et Mitchell, J. (2014). Dropout prediction in MOOC using learners activity features. *eLearning papers*.
- Krumm, A., Waddington, R., Lonn, S. et Teasley, S. (2014). A learning management system-based early warning system for academic advising in undergraduate engineering. *Handbook of Learning Analytics: Methods, Tools and Approaches*.
- Lim, F., O'Halloran, K. et Podlasov, A. (2012). Spatial pedagogy: Mapping meaning in the use of classroom space. *Cambridge Journal of Education*, 42, 2, p. 235-251.
- Lonn, S., Krumm, A., Waddington, R. et Teasley, S. (2012). Bridging the gap from knowledge to action: Putting analytics in the hands of academic advisors. *Proc. LAK*.
- Sauvé, L. & Leclerc, G. & Nadeau J.-R. (1993). Les effets d'autonomisation et de socialisation dans les programmes universitaires individualisés. In *Individualisation de la formation : recherche et exemples*. Publié sous la direction de L. Sauvé et G. Pineau. ISBN 2-7624-4400-4. pp 13-60, 1993.
- Thille, C. et Smith, J. (2011). Cold rolled steel and knowledge: What can higher education learn about productivity? *Change*, 43(2), p. 21-27.

LES AUTEURS

Geoffray Bonnin, Maître de conférences en informatique, Université de Lorraine
 Anne Boyer, Professeur en informatique, Université de Lorraine