

Les parcours étudiants à l'ère des données : modéliser, comprendre et agir sur les dynamiques d'engagement

Résumé :

Les environnements numériques d'apprentissage génèrent une masse croissante de données sur les activités des étudiant.e.s. Ces traces, bien que prometteuses pour l'analyse de l'engagement et de la réussite, demeurent souvent fragmentées et difficilement exploitables. Cet article propose une approche fondée sur le *process mining* pour transformer ces données brutes en informations pédagogiquement pertinentes. En s'appuyant sur un cours universitaire suivi par 182 étudiants en sciences de gestion, l'étude montre comment le *process mining*, sorti de son contexte industriel initial permet de modéliser les trajectoires d'apprentissage réelles et d'identifier des profils dynamiques d'apprenants. L'analyse distingue quatre types de comportements en accord avec la littérature (universitaires/scolaires, engagés tardifs, décrocheurs tardifs et non participants) et révèle une corrélation forte entre la régularité de l'engagement et la performance académique. Ces résultats soulignent le potentiel du *process mining* comme outil de diagnostic et de pilotage pédagogique capable de relier visualisation des parcours et accompagnement individualisé. Au-delà de la preuve de concept, l'article ouvre la voie à une pédagogie conçue comme un processus dynamique, dans laquelle les dispositifs d'enseignement peuvent être ajustés en fonction des pratiques effectives plutôt que des seuls modèles prescrits.

Mots-clés : learning analytic, processus, données

INTRODUCTION

Depuis plus d'une décennie, l'enseignement supérieur connaît une transformation liée à l'apparition, puis à la généralisation, des environnements numériques d'apprentissage (LMS) tels que Moodle ou Blackboard. Ces plateformes captent un volume considérable de données comme le temps de connexion, les séquences de navigation, les résultats aux activités, la participation à celles-ci ou encore les téléchargements de documents. Cette disponibilité de traces d'activité a donné naissance au champ des Learning Analytics (LA) dont l'ambition est de mieux comprendre les parcours étudiants et d'optimiser les dispositifs pédagogiques (Peraya, 2019).

Dans l'enseignement des sciences de gestion, le champ des LA prend une importance encore plus particulière. Les cours de type STEM, dits de support (informatique, mathématiques, statistiques, ...) jouent un rôle essentiel dans la formation. Cependant, leur position périphérique par rapport aux disciplines perçues comme centrales (finances, marketing, ...) rend leur appropriation plus difficile. Les étudiants ont souvent du mal à percevoir la valeur directe de ces compétences techniques, ce qui affecte leur motivation et leur engagement (Glynn et al., 2007). Dans ce contexte, les approches issues des LA apparaissent particulièrement pertinentes. Elles permettent de mieux comprendre les dynamiques d'apprentissage réelles sans tenir compte des à priori contextuels des étudiants.

Les promesses associées aux LA demeurent en partie inachevées. Les données recueillies sont souvent hétérogènes, bruitées et difficiles à exploiter par les enseignants. Ils peinent à en extraire des informations directement mobilisables pour ajuster leurs pratiques (Peraya, 2019). Les traces collectées apparaissent fragmentées et peu significatives en l'état. De nombreux travaux soulignent la nécessité de reconstruire et de modéliser ces traces afin de les rendre interprétables et analytiquement pertinentes (Settouti & al, 2009 ; Peraya & al, 2009). Ce besoin est aujourd'hui d'autant plus crucial que la multiplication des sources de données rend encore plus délicate la constitution de journaux cohérents et exploitables (Ouakrat & Mesangeau, 2016).

L'un des principaux défis des LA contemporains ne réside pas uniquement dans la collecte d'informations mais dans la production d'analyses réellement pertinentes. Dans le cas contraire, la masse d'informations disponibles risque de se transformer en bruit plutôt qu'en levier de pilotage pédagogique. À cela s'ajoute la question des profils d'apprenants mobilisés dans les démarches pédagogiques. Ceux-ci sont encore fréquemment définis selon des catégories a priori ou des indicateurs limités (notes, assiduité, ...). Cela reflète imparfaitement la richesse et l'évolution des comportements d'apprentissage (Papet, 2025).

Cette situation engendre un double paradoxe : les établissements disposent d'une quantité croissante de données sur les étudiants mais leur potentiel pédagogique reste largement sous-exploité. Parallèlement, les apprenants continuent d'être classés dans des profils statiques alors que leurs pratiques et leur engagement fluctuent.

Dans cet optique, le *process mining* (van der Aalst, 2016) constitue une approche prometteuse et innovante. Il est d'abord défini comme une discipline consacrée à l'analyse des processus dans le monde de l'entreprise et des chaînes de production. Il permet d'optimiser les flux de travail, de réduire les coûts et d'améliorer la conformité des produits. Son détournement vers d'autres contextes, comme l'éducation ou la santé, s'avère néanmoins possible. Cette transférabilité tient du fait que, quel que soit le domaine, les traces suivent une logique comparable à celle des processus métiers. En plus de cela, une des composantes du *process mining*, le *trace clustering* (Song et al., 2009), rend possible la construction de profils dynamiques fondés sur les comportements effectifs des étudiants plutôt que sur des catégories théoriques préétablies.

Cette démarche ouvre des perspectives nouvelles pour la personnalisation pédagogique. Les enseignants peuvent visualiser les trajectoires d'apprentissage réellement suivies, détecter précocement les étudiants à risque de décrochage et adapter leurs interventions en conséquence.

Dans cet article, nous proposons d'explorer cette voie en articulant trois contributions :

- La présentation d'un cadre conceptuel intégrant le *process mining* et le *trace clustering* pour l'analyse des parcours étudiants.
- Une application originale de ces méthodes dans le cadre d'une formation universitaire en management et gestion.
- Une discussion sur la pertinence des résultats obtenus et leurs implications pédagogiques réelles.

La suite de l'article est organisée comme suit. La section 2 présente un état de l'art des travaux sur les Learning Analytics, la question des profils d'apprenants et les applications récentes du *process mining* en contexte éducatif. La section 3 propose un cadre conceptuel articulant *process mining* et *trace clustering* pour la construction de profils dynamiques ainsi qu'une discussion sur leur mise en œuvre dans les environnements numériques d'apprentissage. La section 4 présente les résultats et implications de la méthodologie proposée. La section 5 explore, finalement, les implications pédagogiques de cette approche.

ÉTAT DE L'ART

L'ESSOR DES DONNÉES ÉDUCATIVES ET DES *LEARNING ANALYTICS*

La généralisation des environnements numériques d'apprentissage a profondément modifié la manière dont les universités conçoivent, diffusent et évaluent les apprentissages. Chaque interaction étudiante génère désormais une trace. Cet ensemble constitue une source précieuse d'informations pour mieux comprendre l'engagement et les parcours d'apprentissage (Peraya, 2019).

Les Learning Analytics (LA) se sont développées précisément pour tirer parti de ces données. Leur objectif est double : fournir aux enseignants et aux institutions des indicateurs permettant de suivre la progression des étudiants et contribuer à la personnalisation des dispositifs pédagogiques afin de soutenir la réussite (El Kechai, s.d).

Plusieurs auteurs mettent en évidence le potentiel des Learning Analytics pour transformer les pratiques pédagogiques. Peraya (2019) souligne que les LA ne se limitent pas à une fonction de reporting mais permettent d'observer finement les comportements d'apprentissage et Cherigny & al (2021) insiste sur le potentiel des LA pour la conception de parcours adaptatifs où les contenus et les activités proposés aux étudiants peuvent être ajustés en fonction de leurs besoins, de leurs performances et de leur niveau d'engagement.

Cependant, l'exploitation de ces données reste un défi. Les traces sont souvent hétérogènes, distribuées sur plusieurs systèmes et difficiles à consolider. Leur qualité varie en fonction des contextes d'enseignement et des outils utilisés, ce qui complique leur interprétation et limite leur potentiel prédictif (Sacré & al, 2020).

DES TRACES BRUTES À LA MODÉLISATION DES PARCOURS

Les données captées par les plateformes ne sont que rarement prêtes à l'emploi. Settouti & al (2009) les qualifient d' « inscriptions » et souligne leur caractère brut et fragmenté. Pour être utiles à la recherche ou à la pratique pédagogique, ces traces doivent être reconstruites, agrégées et parfois enrichies afin de leur donner une signification analytique. Peraya & al (2009) insistent sur la nécessité d'opérer ce travail à différents niveaux de granularité selon l'objectif poursuivi.

Sassi et Laroussi (2012) ainsi que Ouakrat et Mesangeau (2016) ont proposé des cadres méthodologiques pour structurer les traces d'apprentissage et les transformer en données exploitables. Cependant, avec la diversification des outils et la complexité croissante des parcours d'études, la constitution de journaux cohérents et comparables demeure un problème majeur (Djouad, 2011). Sans ce travail de mise en forme, la

masse de données produites peut générer une surcharge informationnelle plutôt qu'un soutien efficace à la décision pédagogique.

PROFILS D'APPRENANTS : ENTRE TYPOLOGIES FIGÉES ET DYNAMIQUES

L'un des usages récurrents des Learning Analytics consiste à classifier les apprenants en profils afin de mieux cibler les interventions pédagogiques. Traditionnellement, ces profils sont définis sur la base de variables simples comme l'assiduité, les notes ou la participation aux activités en ligne. Cette approche permet de repérer les étudiants « en difficulté » ou « performants » mais elle tend à figer des comportements qui, en réalité, évoluent tout au long du semestre.

Papet (2025) a mis en évidence l'existence de profils académiques distincts à l'université et souligne leur caractère évolutif ainsi que l'importance de suivre les étudiants de manière longitudinale pour saisir leurs transitions entre phases d'engagement et de désengagement. Dans le même sens, Cisel (2017) met en avant l'idée de profils dynamiques, construits à partir de données d'interaction et régulièrement mis à jour afin de refléter la réalité des apprentissages.

LE PROCESS MINING ET LE TRACE CLUSTERING EN CONTEXTE ÉDUCATIF

Le *process mining* apparaît comme une approche particulièrement adaptée pour analyser les parcours d'apprentissage. Issu du domaine de la science des processus (van der Aalst, 2016), il permet de reconstruire automatiquement, à partir de journaux d'événements, les séquences réelles d'activités suivies par les étudiants. Plusieurs travaux récents ont démontré son potentiel en contexte éducatif, notamment pour visualiser les chemins d'apprentissage, identifier des comportements inattendus ou détecter des écarts par rapport aux scénarios pédagogiques prévus (Bogarin & al, 2018).

Pour améliorer la lisibilité et la pertinence des modèles générés, le trace clustering est souvent utilisé en amont afin de regrouper les apprenants présentant des comportements similaires avant de réaliser la découverte de processus. Seok & al. (2023) montrent que cette approche améliore la précision des modèles et facilite leur interprétation. En contexte éducatif, elle offre également la possibilité d'identifier des profils comportementaux fondés sur les traces effectives plutôt que sur des catégories préétablies.

Limites et besoins de recherche

Malgré les avancées, plusieurs défis demeurent. La question de l'interopérabilité des données entre plateformes reste ouverte, tout comme celle de leur actionnabilité pour les enseignants qui doivent pouvoir interpréter les résultats sans expertise technique avancée. L'exploitation des traces soulève des enjeux de gouvernance et d'éthique tels que le consentement des étudiants, le respect de la vie privée et la prévention de la stigmatisation.

CADRE CONCEPTUEL ET MÉTHODOLOGIE PROPOSÉE

PRINCIPES MÉTHODOLOGIQUES

L'approche proposée repose sur un enchaînement d'étapes visant à transformer les traces brutes issues des plateformes d'apprentissage en profils d'apprenants. Chacune de ces étapes est illustrée ci-dessous par un exemple basé sur un jeu de données simulé.

Étape 1 : Collecte et préparation des données

La première étape consiste à extraire les journaux d'événements depuis les environnements numériques d'apprentissage (Moodle, Blackboard, ...). Ces journaux contiennent généralement l'identifiant de l'étudiant, l'activité réalisée, l'horodatage et la ressource associée. Un prétraitement est indispensable pour en garantir la qualité comme la suppression des doublons et exclusion des données parasites. Cette phase aboutit à un *log* d'événements structuré, compatible avec les outils de *process mining* et prêt pour l'analyse.

Exemple :

Cas	Activité	Horodatage
Clara	Lecture_Chapitre_1	2025-10-01 09:05
Clara	Quiz_1	2025-10-01 09:45
Clara	Forum_Question_1	2025-10-01 10:00
Julien	Lecture_Chapitre_1	2025-10-01 08:55
Julien	Quiz_1	2025-10-01 09:30
Julien	Forum_Question_1	2025-10-01 09:45
Sarah	Lecture_Chapitre_1	2025-10-01 10:15
Sarah	Quiz_1	2025-10-01 10:50
Sarah	Forum_Question_1	2025-10-01 11:00
Pierre	Quiz_1	2025-10-01 09:20
Lucas	Quiz_1	2025-10-01 09:25
Lucas	Abandon_Cours	2025-10-02 08:00
Amine	Quiz_1	2025-10-01 08:50
Lisa	Quiz_1	2025-10-01 09:00
Lisa	Abandon_Cours	2025-10-08 09:10

Étape 2 : Découverte de processus (process mining)

Une fois les données préparées, la deuxième étape consiste à reconstruire les parcours d'apprentissage par découverte de processus. L'utilisation d'algorithmes tels que l'*Inductive Miner* ou le *Fuzzy Miner* (Leemans & al, 2014) permet de modéliser visuellement les séquences d'activités réellement suivies par les étudiants. Cette analyse révèle non seulement le chemin dominant (parcours le plus fréquent) mais également les comportements divergents, les points de sortie prématurée et les goulots

d'étranglement dans le parcours. Elle constitue un outil puissant pour comparer le scénario pédagogique prévu et les comportements observés et pour identifier les écarts qui peuvent nécessiter un ajustement du dispositif.

Exemple :

- 3/7 des étudiants suivent le parcours prévu (Lecture → Quiz → Forum)
- 4/7 passent directement au quiz
- 2/7 abandonnent le cours après avoir répondu au quizz

Étape 3 : Segmentation par clustering

La troisième étape applique le *trace clustering* qui regroupe les étudiants en sous-populations homogènes selon leurs comportements. Cette segmentation réduit la complexité des modèles de processus et facilite leur interprétation par les enseignants. Elle permet de mettre en évidence des profils.

Exemple :

- Cluster 1 – Scolaires (3/7) : suivent le parcours prévu en consultant les ressources, réalisant le quiz et participant au forum.
- Cluster 2 – Énigmatiques (2/7) : passent directement au quiz avant de poursuivre le parcours.
- Cluster 3 – Résignés (2/7) : réalisent le quiz puis interrompent leur progression et n'accomplissent plus d'activités par la suite.

Étape 4 : Construction de profils dynamiques

Ces clusters sont ensuite traduits en profils pédagogiques dynamiques. L'une des forces de l'approche proposée réside dans la possibilité de mettre à jour régulièrement ces profils de sorte qu'un étudiant puisse changer de catégorie au fil du semestre. Cette dynamique rend l'analyse plus pertinente que les typologies figées en reflétant l'évolution des comportements d'apprentissage et en permettant un suivi longitudinal des trajectoires.

Étape 5 : Restitution et action pédagogique

La dernière étape consiste à restituer les résultats de manière intelligible et actionnable. Les informations issues du process mining et du clustering peuvent être intégrées dans des tableaux de bord accessibles aux enseignants, accompagnées d'indicateurs synthétiques (proportion d'étudiants dans chaque profil, progression dans le temps). Cette restitution facilite la mise en place d'actions ciblées : accompagnement personnalisé pour les étudiants en difficulté, enrichissement du contenu pour les apprenants avancés, ou encore ajustement de la séquence pédagogique pour améliorer l'engagement global.

Exemple : Après la mise en place d'une séance de questions-réponses, la proportion d'étudiants du Cluster 3 – Résignés est passée de 2 sur 7 à 1 sur 12, montrant une réduction notable du décrochage.

RÉSULTATS

Les données empiriques mobilisées dans cette étude ont été collectées auprès de 182 étudiants inscrits en bachelier en sciences de gestion, en sciences économiques et de gestion et en ingénieur de gestion. Elles concernent un cours de transformation numérique dont la partie pratique est consacrée à l'initiation à la programmation en Python.

Le dispositif pédagogique comprenait six séances de travaux pratiques. Chacune de ces séances étant associée à une activité d'évaluation formative sous la forme de quiz récapitulatifs disponibles sur la plateforme en ligne. Ces activités visaient à consolider les apprentissages et à suivre la progression des étudiants au fil du quadrimestre.

Le jeu de données obtenu couvre l'ensemble du quadrimestre et comporte les éléments suivants :

- L'identifiant de l'étudiant
- L'activité réalisée (séance correspondante)
- Le statut de réalisation (activité non effectuée, effectuée avec succès ou effectuée sans succès),
- La note obtenue à l'activité
- La note finale attribuée au cours a été aussi ajoutée

Ces informations constituent un journal d'événements structuré qui permettant de reconstituer les parcours effectifs des étudiants et d'analyser les liens entre participation aux activités pratiques et performance académique globale.

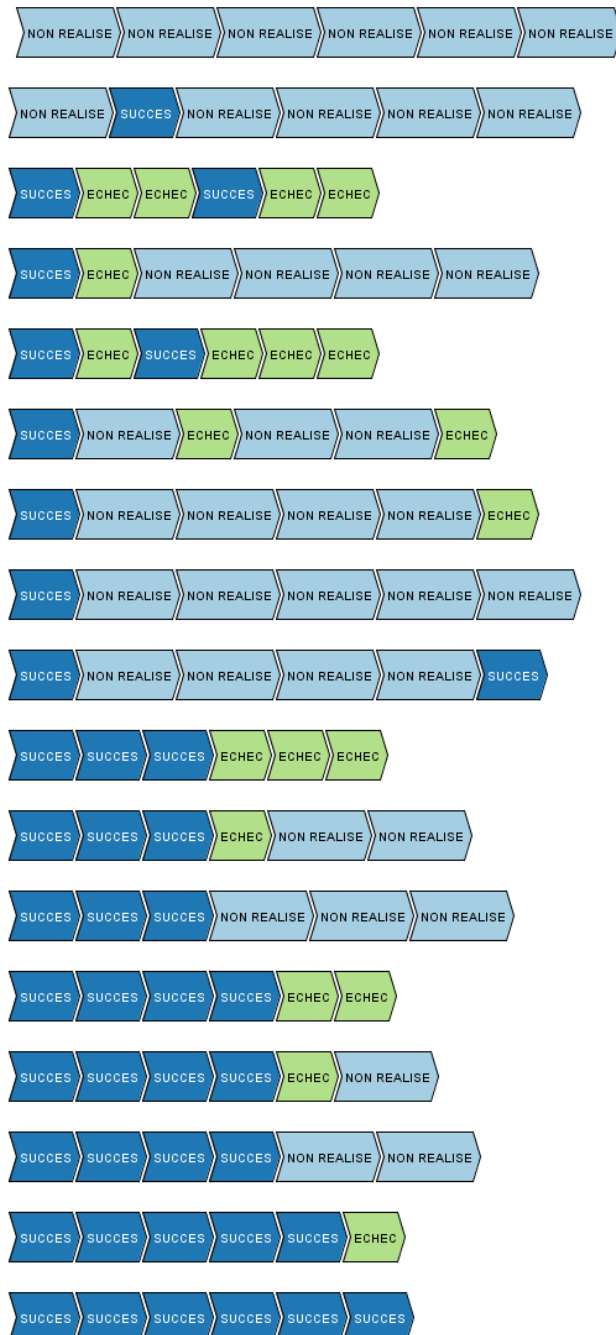
PARCOURS ATEMPORELS DES ÉTUDIANTS

Le choix d'une représentation atemporelle vise à simplifier la lecture en se concentrant sur la combinaison des résultats plutôt que sur l'ordre réel des activités. Cette approche permet de comparer plus aisément les trajectoires similaires entre étudiants et de mettre en évidence la structure générale des parcours. Cette représentation ne tient pas compte du fait que certains aient travaillé de façon régulière, de manière aléatoire ou uniquement à l'approche de l'évaluation finale. La première case représente donc l'activité 1 pour l'ensemble des étudiants, la deuxième case l'activité 2, la troisième case l'activité 3, ...

La figure ci-dessous illustre les trajectoires suivies sur les six activités. Chaque ligne correspond à un cheminement type observé parmi les

étudiants. On constate une hétérogénéité dans les pratiques (17 profils pour 182 étudiants) mais avec une certaine logique inhérente. La combinaison du premier profil et des deux derniers représente 89,6% de l'ensemble des étudiants, c'est-à-dire ceux qui ont réussi l'ensemble des activités (ou presque) et ceux qui n'ont pas essayés.

Figure 1 : profil atemporel des étudiants



Un constat qui peut être également observé est que la non-réalisation ou l'échec d'une activité entraîne dans presque l'ensemble des cas (92% des cas où cela se produit) un décrochage des activités suivantes.

Figure 2 : modélisation du parcours atemporel des étudiants



La figure ci-dessus illustre l'intérêt du process mining pour l'analyse des parcours étudiants. En un seul diagramme et avec un seul outil, il restitue une vision claire et synthétique de la situation, ici pour les activités 2, 3, 4 et 5. Trois comportements apparaissent immédiatement : certains étudiants tentent l'activité et réussissent (branche inférieure), d'autres l'essayent mais échouent (branche du milieu), tandis qu'un troisième groupe (branche supérieure) ne s'y engage pas du tout. L'apport essentiel de l'approche réside dans sa capacité à quantifier ces comportements et à visualiser directement leur répartition au sein de la cohorte. On obtient ainsi, en quelques instants, une lecture structurée de l'engagement étudiant qui serait difficile à dégager à partir des seules données brutes.

Les différences, mineures, qui apparaissent (par exemple un passage de 4 à 2 étudiants pour la branche inférieure de la première partie de la figure ou un passage de 3 à 5 pour la branche centrale de la deuxième partie de la figure) sont des cas différents que l'on peut investiguer en modifiant le niveau de granularité de l'outil pour une analyse plus profonde mais non nécessaire à la compréhension globale des dynamiques.

PARCOURS TEMPORELS DES ÉTUDIANTS

La figure 3 présente la répartition temporelle des activités réalisées par les étudiants au cours du quadrimestre. L'axe horizontal représente le temps, couvrant l'ensemble de la période académique, tandis que l'axe vertical (Y) correspond aux étudiants, ordonnés en fonction de leur date de première activité. Les étudiants situés en haut du graphique sont ceux qui ont commencé le plus tôt et ceux en bas sont ceux qui ont attendu le plus longtemps avant d'entamer leurs activités. Chaque point matérialise une activité effectuée à une date précise par un étudiant donné.

L'observation du graphique permet de distinguer quatre zones principales. La première zone, en haut du graphique, est caractérisée par une forte

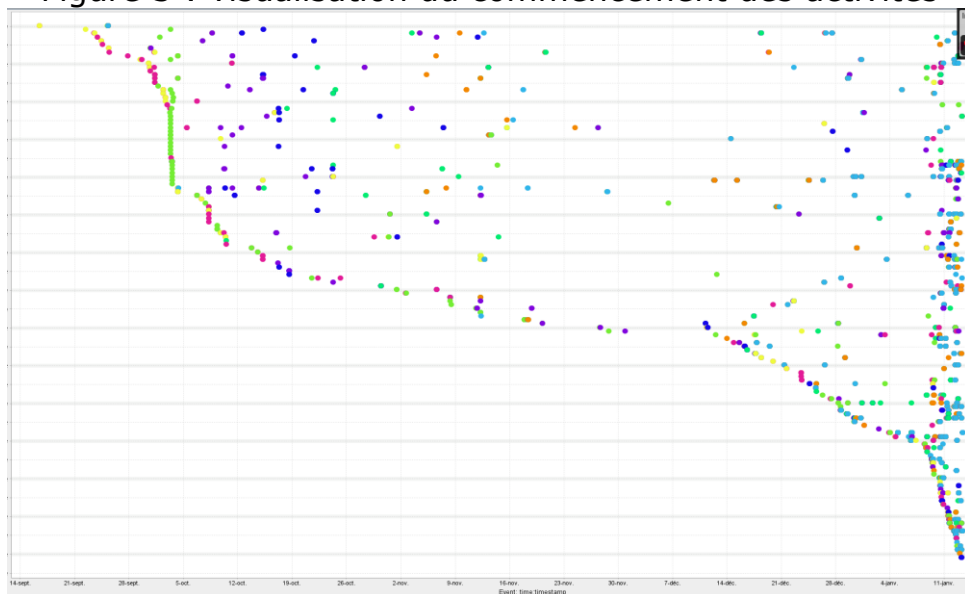
concentration verticale de points au début du quadrimestre. Elle correspond aux étudiants qui se sont engagés immédiatement dans les activités dès leur mise à disposition.

La deuxième zone, située en dessous, regroupe des étudiants dont la progression est moins structurée. Leurs activités apparaissent de manière plus dispersée et aléatoire dans le temps.

La troisième zone, plus verticale et apparaissant vers la fin des cours, reflète un regain d'activité lié au début de la période de révision. Les étudiants concernés tendent à concentrer leurs efforts dans un laps de temps réduit et juste avant les évaluations finales.

La quatrième zone, très marquée et verticale en bas du graphique, illustre le comportement des étudiants qui n'ont commencé les activités que dans les jours précédant directement l'examen. Ce découpage en quatre zones met en évidence des profils d'engagement différenciés allant de la régularité précoce à la procrastination maximale. Ces dynamiques temporelles constituent un indicateur précieux du rapport des étudiants à l'organisation de leur apprentissage et peuvent être utilisées pour identifier des besoins pédagogiques différenciés.

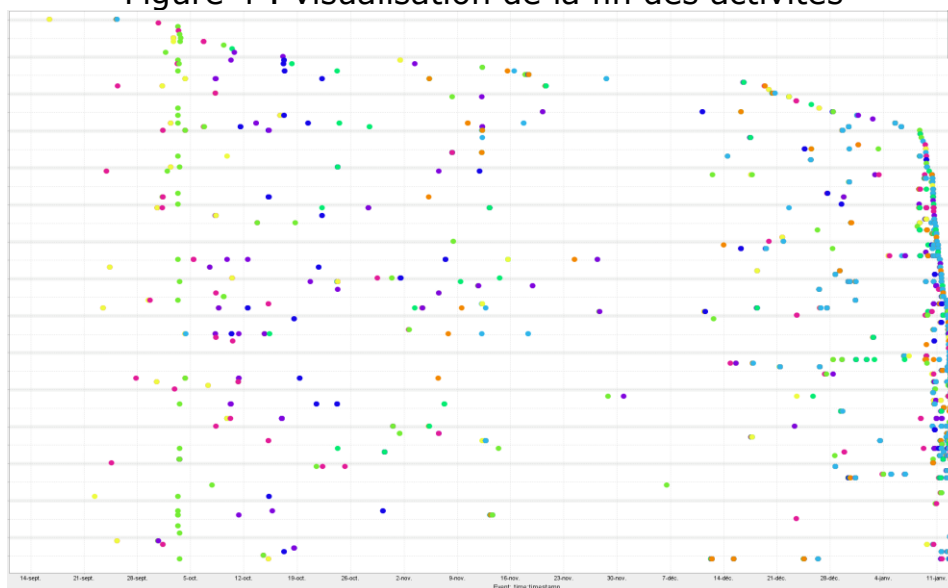
Figure 3 : visualisation du commencement des activités



La figure 4 adopte un point de vue opposé à celui de la figure précédente en représentant les étudiants classés non plus selon leur date de première activité mais selon leur date de dernière activité effectuée. Ainsi, les étudiants situés en haut du graphique sont ceux qui ont terminé leurs activités le plus tôt tandis que ceux en bas sont ceux qui ont poursuivi jusqu'à la fin du quadrimestre.

On observe que la partie supérieure du graphique (environ un cinquième de la hauteur totale) regroupe les étudiants qui n'ont pas mené leurs activités jusqu'au bout. Leur dernière activité étant intervenue bien avant la fin de la période. La partie inférieure, en revanche, ne permet pas de distinguer clairement les profils d'engagement. Elle rassemble à la fois les étudiants qui ont suivi les activités de manière régulière et les ont achevées en fin de quadrimestre et que ceux qui se sont concentrés sur les activités à la dernière minute mais ont néanmoins terminé à temps.

Figure 4 : visualisation de la fin des activités



DISCUSSION ET IMPLICATIONS PÉDAGOGIQUES

L'analyse des parcours étudiants, qu'elle soit menée sous une perspective atemporelle (c'est-à-dire en considérant les activités sans prendre en compte leur séquençement temporel) ou temporelle (en analysant la dynamique et la régularité de l'engagement), conduit principalement à deux constats généraux. D'une part, les trajectoires d'apprentissage ne sont pas homogènes : elles traduisent une diversité de comportements et de stratégies qui ne peuvent être réduites à une simple opposition entre étudiant·es « présents » et « absents ». D'autre part, ces stratégies d'engagement s'avèrent fortement corrélées aux résultats finaux. Cela suggère que la manière dont les étudiant·es interagissent avec la plateforme et les dispositifs d'apprentissage est aussi déterminante que l'effort total consenti. Sur cette base, il est possible de distinguer quatre grands profils d'étudiants.

LES NON PARTICIPANTS.

Cette première catégorie rassemble celles et ceux qui ne réalisent pas les activités d'apprentissage proposées sur la plateforme et qui, dans la majorité des cas, échouent à l'examen final. Ce groupe représente une

part significative de la cohorte et illustre le phénomène du décrochage complet. Toutefois, l'analyse fine des résultats nuance ce constat. Une proportion d'étudiants de cette catégorie parvient malgré tout à réussir l'examen. Cette réussite peut être interprétée comme le signe d'un désengagement vis-à-vis du dispositif numérique plutôt que d'un décrochage total du cours. En d'autres termes, ces étudiant·es auraient choisi d'autres modes d'apprentissage (tels que l'étude autonome à partir de supports papier ou encore le travail collaboratif de synthèse avec leurs pairs) qui leur permettent d'atteindre l'objectif final sans recourir à la plateforme. Cette distinction entre décrochage numérique et décrochage académique est essentielle et elle invite à ne pas assimiler l'absence d'interactions avec les outils numériques à un échec inéluctable mais à envisager une pluralité de stratégies d'appropriation du cours.

LES UNIVERSITAIRES/SCOLAIRES.

La deuxième catégorie correspond aux étudiant·es décrits par Papet (2025) sous le terme d'« universitaires/scolaires ». Ils représentent environ 10 % de la cohorte et constituent le groupe le plus performant. Leur parcours est marqué par une participation régulière et synchrone aux activités proposées. Les tests formatifs sont réalisés au fur et à mesure de l'avancement du cours, validés dès la première tentative et intégrés dans une logique de consolidation immédiate des apprentissages. Cette approche traduit une articulation réussie entre l'assiduité aux enseignements présentiels et l'exploitation active des ressources numériques. Ces étudiant·es incarnent une forme de « bonne pratique » et leur réussite à l'examen final confirme la pertinence de cette stratégie.

LES ENGAGÉS TARDIFS.

Un troisième profil correspond aux étudiant·es qui concentrent leurs efforts dans une période limitée, généralement juste avant l'échéance de l'examen. Leur comportement se caractérise par une participation asynchrone. Les activités proposées en cours de semestre ne sont pas mobilisées en temps réel mais utilisées comme outils de révision intensive en amont de l'évaluation certificative. Ce groupe constitue une part importante de la cohorte (57%) et représente une stratégie répandue parmi les étudiants·es. Les résultats de cette catégorie sont contrastés : le taux de réussite est équilibré mais inférieur à celui observé chez les universitaires/scolaires. Autrement dit, cette stratégie « de dernière minute » permet à certain·es de réussir mais elle ne constitue pas une garantie d'efficacité pédagogique. Elle suggère que la temporalité de l'engagement (c'est-à-dire la régularité et la répartition des efforts dans la durée) joue un rôle déterminant dans l'appropriation en profondeur des connaissances.

LES DÉCROCHEURS TARDIFS.

La dernière catégorie regroupe les étudiant·es qui débutent leur parcours avec une dynamique proche de celle des universitaires/scolaires (assiduité aux activités, réussite aux premiers tests formatifs) mais qui interrompent progressivement leur engagement en cours de semestre. Ces « décrocheurs tardifs » apparaissent comme la catégorie la plus fragile puisqu'ils affichent le taux d'échec le plus élevé. Leur trajectoire met en lumière le rôle crucial important. Un engagement initial soutenu ne suffit pas à garantir la réussite si celui-ci n'est pas maintenu jusqu'à l'évaluation finale. Cette catégorie interroge particulièrement les dispositifs d'accompagnement puisqu'elle semble concerner des étudiant·es motivés et capables mais confrontés à des obstacles (organisationnels, motivationnels ou contextuels) qui entraînent un décrochage progressif.

L'identification de ces quatre profils permet de mieux comprendre les menant à la réussite ou à l'échec. Elle confirme également que l'échec n'est pas réductible à un manque de capacité cognitive mais qu'il est étroitement lié à la régularité de l'engagement.

Qu'elle soit appréhendée de manière atemporelle ou temporelle, l'analyse met en évidence une même réalité, la diversité des trajectoires étudiantes et leur lien direct avec la performance académique. Ces résultats invitent à envisager des dispositifs pédagogiques différenciés, capables de soutenir les étudiant·es réguliers dans leur dynamique vertueuse, d'accompagner ceux qui privilégient une préparation tardive et surtout de repérer précocement les risques de décrochage afin de mettre en place des mesures de remédiation adaptées.

L'ensemble de ces résultats met en évidence la pertinence du recours au *process mining* pour analyser les parcours étudiants. La méthode s'est révélée particulièrement efficace dans ce cas précis car elle rend immédiatement visibles des comportements qui resteraient difficiles à appréhender par des approches statistiques classiques. Elle permet de distinguer en un seul diagramme des trajectoires contrastées (étudiants réguliers, engagés tardifs, décrocheurs ou non participants) et d'en relier directement l'observation aux résultats académiques. Sa capacité à combiner une lecture atemporelle (centrée sur la nature et la fréquence des activités) et une lecture temporelle (attentive à la dynamique et à la régularité de l'engagement) constitue un atout majeur pour saisir la complexité des pratiques étudiantes.

Au-delà de la mise en évidence de profils, le *process mining* offre une visualisation claire et synthétique qui facilite l'interprétation et le transfert des résultats vers les acteurs pédagogiques. Il ne s'agit donc pas seulement d'un outil de recherche mais également d'un instrument de

diagnostic et de pilotage utile pour repérer rapidement les situations de décrochage et envisager des actions de soutien adaptées.

Dans ce cas précis, il a permis de décrire la diversité des trajectoires étudiantes et de mettre en lumière le lien entre engagement, temporalité de l'investissement et réussite académique, confirmant ainsi sa valeur ajoutée comme méthode d'analyse en sciences de l'éducation et en learning analytics.

CONCLUSION

La réflexion développée dans ce travail met en évidence l'intérêt d'aborder les dispositifs d'enseignement sous l'angle des processus en se basant sur des méthodes d'analyse issues du monde de l'entreprise. Concevoir l'enseignement comme un processus permet de dépasser une vision strictement transmissive du savoir pour y introduire une logique d'optimisation et d'amélioration. Cette perspective ouvre un champ de recherche particulièrement qui est celui d'une pédagogie conçue comme un système dynamique.

Cette analogie entre enseignement et production ne vise pas à réduire l'enseignement à une simple chaîne d'opérations mécaniques. Tout comme une entreprise cherche à améliorer la qualité de ses produits et la satisfaction de ses clients, l'institution d'enseignement peut mobiliser ces outils pour analyser la qualité des apprentissages, repérer les points de friction et ajuster les dispositifs de formation. Une telle approche s'inscrit dans la continuité des démarches qualité mais elle en renouvelle les fondements méthodologiques.

Le recours à des outils issus du *process mining* en support des *Learning Analytics* illustre concrètement cela. En exploitant les traces d'activité laissées par les étudiants sur les plateformes numériques, il devient possible de modéliser les processus d'apprentissage réels et de les confronter aux processus prescrits. Cette confrontation permet non seulement d'identifier les écarts, mais aussi de comprendre les stratégies d'adaptation mises en œuvre par les apprenants. Dès lors, la modélisation de l'enseignement devient un instrument de pilotage et d'aide à la décision.

Les résultats obtenus confirment la pertinence de cette approche. L'analyse des données recueillies a permis de mettre en évidence plusieurs régularités dans les comportements d'apprentissage et de mieux comprendre la manière dont les étudiants interagissent avec les dispositifs pédagogiques. Les modèles issus du *process mining* ont notamment révélé des chemins d'apprentissage récurrents, des phases de décrochage caractéristiques ainsi qu'une forte influence de la temporalité des activités. Ces observations valident l'idée selon laquelle les processus éducatifs

peuvent être modélisés de manière rigoureuse. Elles ouvrent ainsi la voie à des dispositifs de formation plus adaptatifs qui sont capables d'évoluer en fonction des pratiques réelles plutôt que des seules intentions pédagogiques.

BIBLIOGRAPHIE

Bogarín, A., Cerezo, R., & Romero, C. (2018). *A survey on educational process mining. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(1), e1230.

Cherigny, F., El Kechai, H., Iksal, S., Lefevre, M., Labarthe, H., et al. (2020). *L'analytique des apprentissages avec le numérique*. Direction du numérique pour l'éducation (DNE-TN2). <https://hal.archives-ouvertes.fr/hal-02912386>

Djouad, T. (2011). *Ingénierie des indicateurs d'activités à partir de traces modélisées pour un Environnement Informatique d'Apprentissage Humain*. Thèse de doctorat, Université Claude Bernard-Lyon I & Université Mentouri-Constantine.

El Kechai, H. (s.d.). *Learning Analytics : utilisation à des fins scolaires et usages* — *Bulletin de veille* n° 2.

Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2007). Nonscience majors learning science: A theoretical model of motivation. *Journal of Research in Science Teaching*, 44(8), 1088-1107.

Leemans, S. J., Fahland, D., & Van Der Aalst, W. M. (2014). Process and deviation exploration with inductive visual miner. In *12th International Conference on Business Process Management, BPM 2014* (pp. 46-50). CEUR-WS. org.

Ouakrat, A., & Mésangeau, J. (2016). Resocialiser les traces d'activités numériques : une proposition qualitative pour les SIC. *Revue française des sciences de l'information et de la communication*, (8). <https://doi.org/10.4000/rfsic.1795>

Papet, J. (2025). Des profils académiques d'étudiants différents à l'Université supérieur : quelles caractéristiques psycho-cognitives spécifiques ? *Revue internationale de pédagogie de l'enseignement*, 41(2). <https://doi.org/10.4000/14kwe>

Peraya, D. (2019). Les Learning Analytics en question : Panorama, limites, enjeux et visions d'avenir. *Distances et médiations des savoirs*, 26. <https://doi.org/10.4000/dms.3485>

Peraya, D., Batier, C., Paquelin, D., Rizza, C., & Vieira, M. (2009). Les traces d'usage et l'usage des traces : Le rôle des traces dans l'orientation stratégique des unités de développement de l'eLearning et des dispositifs hybrides dans l'enseignement supérieur. *Le numérique dans l'enseignement et la formation. Analyses, traces et usages*, 37-80.

Sacré, M., Toczek, M.-C., Policard, F., Serres, G., Paulet, C., Petit, S., Neuville, E., & La Fontaine, D. (2020). L'efficacité d'un dispositif d'enseignement hybride en fonction des caractéristiques des étudiants. *Revue internationale des technologies en pédagogie universitaire*, 17(2), 9-29. <https://doi.org/10.18162/ritpu-2020-v17n2-03>

Settouti, L. S., Prie, Y., Marty, J.-C., & Mille, A. (2009). A Trace-Based System for Technology-Enhanced Learning Systems Personalisation. *2009 Ninth IEEE International Conference on Advanced Learning Technologies*, Riga, Latvia, 93-97. <https://doi.org/10.1109/ICALT.2009.159>

Song, M., Günther, C. W., & Van der Aalst, W. M. (2008, September). Trace clustering in process mining. In *International Conference on Business Process Management* (pp. 109-120). Berlin, Heidelberg: Springer.

Song, M., Günther, C. W., & van der Aalst, W. M. P. (2009). Trace Clustering in Process Mining. In Ardagna, D., Mecella, M., & Yang, J. (Eds.), *Business Process Management Workshops. BPM 2008. Lecture Notes in Business Information Processing*, vol. 17. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-00328-8_11

van der Aalst, W. M. P. (2016). *Process Mining: Data Science in Action* (2nd ed.). Springer.