THÈSE DE DOCTORAT

DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE-FRANCHE-COMTÉ

PRÉPARÉE À L'UNIVERSITÉ DE FRANCHE-COMTÉ

École doctorale n°37

Sciences Pour l'Ingénieur et Microtechniques

Doctorat d'Informatique

par

DANIEL ANDERSON SOTO FORERO

Adaptation en temps réel d'une séance d'entraînement par intelligence artificielle

Thèse présentée et soutenue à Besançon, le 17 septembre 2025

Composition du Jury :

HULK INCROYABLE	Professeur à l'Université de Gotham City Commentaire secondaire	Président
AMERICA CAPTAIN	Professeur à l'Université USA	Rapporteur
MAN SUPER	Professeur à l'Université de Gotham City	Examinateur
M. MAN BAT	Professeur à l'Université de Gotham City	Directeur de thèse
M. VOLWERINE THE	Professeur à l'Université de Gotham City	Codirecteur de thèse
Mme MAN PAC	Professeur quelque part	Invité

N° | X | X | X |

REMERCIEMENTS

SOMMAIRE

I

CONTEXTE ET PROBLÉMATIQUES

1

INTRODUCTION

Ce chapitre introductif vise à expliquer les termes du sujet et à introduire la thématique principale abordée. Il présente la problématique de cette thèse, introduit les différentes contributions proposées et annonce le plan du manuscrit.

Comme l'indique [Nkambou et al., 2010], les environnements informatiques pour l'apprentissage humain (EIAH) sont des outils proposant des services devant permettre aux apprenants d'acquérir des connaissances et de développer des compétences dans un domaine spécifique. Pour fournir des services efficaces, le système doit intégrer une représentation des connaissances du domaine et des mécanismes pour utiliser ces connaissances. Il doit également être en mesure de raisonner et de résoudre des problèmes.

Le système *Artificial Intelligence - Virtual Trainer* (AI-VT) est un EIAH générique dévéloppé au département d'informatique des systèmes complexes (DISC) de l'institut de recherche FEMTO-ST. Cet outil informatique propose un ensemble d'exercices aux apprenants dans le cadre de séances d'entrainement. AI-VT intègre le fait qu'une séance d'entraînement se situe dans un cycle de plusieurs séances. Les réponses apportées par l'apprenant à chaque exercice sont évaluées numériquement sur une échelle prédéfinie, ce qui permet d'estimer les progrès de l'apprenant et de déduire les sous-domaines dans lesquels il peut avoir des difficultés. Une séance est générée par un système multi-agents associé à un système de raisonnement à partir de cas (RàPC) [Henriet et al., 2017]. Un apprenant choisit le domaine dans lequel il souhaite s'entraîner et AI-VT lui propose un test préliminaire. Les résultats obtenus permettent de placer l'apprenant dans le niveau de maîtrise adéquate. Le système génère ensuite une séance adaptée veillant à l'équilibre entre l'entraînement, l'apprentissage et la découverte de nouvelles connaissances. L'actualisation du niveau de l'apprenant est effectuée à la fin de chaque séance. De cette façon l'apprenant peut avancer dans l'acquisition des connaissances ou s'entraîner sur des connaissances déjà apprises.

Un certain nombre d'EIAH utilisent des algorithmes d'intelligence artificielle (IA) pour détecter les faiblesses et aussi pour s'adapter à chaque apprenant. Les algorithmes et modèles de certains de ces systèmes seront analysés dans les chapitres 3 et 4. Ces chapitres présenteront leurs propriétés, leurs avantages et leurs limites.

Le système AI-VT initial était capable d'ajuster les paramètres de personnalisation d'une séance à l'autre, mais il ne pouvait pas modifier une séance en cours même si certains exercices de celle-ci était trop simples ou trop complexes. Chaque séance était figée et devait être déroulée jusqu'à son terme avant de pouvoir identifier des acquis et des lacunes. Les travaux de cette thèse ont eu pour objectif de pallier ce manque.

La problématique principale de cette thèse est la personnalisation en temps réel du parcours d'apprentissage des apprenants dans le système AI-VT (*Artificial Intelligence - Virtual Trainer*)

Ici *le temps réel* est considéré comme étant le moment où se déroule la séance que l'apprenant est en train de suivre. Par conséquent, l'objectif est de rendre AI-VT plus dynamique pour l'identification des difficultés et l'adaptation du contenu personnalisé en fonction des connaissances démontrées.

La partie suivante présente une liste des principales contributions apportées par cette thèse à la problématique générale énoncée plus haut.

Ce travail de thèse a été éffectué au sein l'Université de Franche-Comté (UFC) devenue depuis le 1er janvier 2025 l'Université Marie et Louis Pasteur (UMLP). Ces recherches ont été menées au sein de l'équipe DEODIS du département d'informatique des systèmes complexes de l'institut de recherche FEMTO-ST, unité mixte de recherche (UMR) 6174 du centre national de la recherche scientifique (CNRS).

1.1/ CONTRIBUTIONS PRINCIPALES

La problématique principale de ces travaux de recherche a été déclinée en plusieurs sous-parties. Ces dernières sont présentées ci-dessous sous la forme de questions. Pour chacune d'elles, une ou plusieurs propositions ont été faites. Voici les questions de recherche abordées et les contributions apportées :

1. **Comment permettre au système AI-VT d'évoluer et d'intégrer de multiples outils ?** Pour répondre à cette question, une architecture modulaire autour du moteur initial d'AI-VT a été conçue et implémentée (ce moteur initial étant le système de raisonnement à partir de cas (RàPC) développé avant le démarrage de ces travaux de thèse).

2. **Comment déceler qu'un exercice est plus ou moins adapté aux besoins de l'apprenant ?** Choisir les exercices les plus adaptés nécessite d'associer une valeur liée à son utilité pour cet apprenant. Les modèles de régression permettent d'interpoler une telle valeur. Pour répondre à cette question, nous avons proposé de nouveaux modèles de régression fondés sur le paradigme du raisonnement à partir de cas et celui des systèmes multi-agents.
3. **Quel modèle et quel type d'algorithmes peuvent être utilisés pour recommander un parcours personnalisé aux apprenants ?** Pour apporter une réponse à cette question, un système de recommandation fondé sur l'apprentissage par renforcement a été conçu. L'objectif de ces travaux est de proposer un module permettant de recommander des exercices aux apprenants en fonction des connaissances demontrées et en se fondant sur les réponses apportées aux exercices précédents de la séance en cours. Ce module de révision de la séance en cours du modèle de RàPC est fondé sur un modèle Bayésien.
4. **Comment consolider les acquis de manière automatique ?** Une séance doit non seulement intégrer des exercices de niveaux variés mais également permettre à l'apprenant de renforcer ses connaissances. Dans cette optique, notre modèle Bayésien a été enrichi d'un processus de Hawkes incluant une fonction d'oubli.

1.2/ PLAN DE LA THÈSE

Ce manuscrit est scindé en deux grandes parties. La première partie contient trois chapitres et la seconde en contient quatre. Le premier chapitre de la première partie (chapitre 2) vise à introduire le sujet, les concepts, les algorithmes associés, présenter le contexte général et l'application AI-VT initiale. Le deuxième chapitre de cette partie présente différents travaux emblématiques menés dans le domaine des environnements informatiques pour l'apprentissage humain. Le chapitre suivant conclut cette première partie en présentant le raisonnement à partir de cas.

Dans la seconde partie du manuscrit, le chapitre 5 explicite l'architecture proposée pour integrer les modules développés autour du système AI-VT initial et élargir ses fonctionalités. Le chapitre 6 propose deux outils originaux fondés sur le raisonnement à partir de cas et les systèmes multi-agents pour résoudre des problèmes de régression de façon générique. Le chapitre 7 présente l'application de ces nouveaux outils de régression dans un système de recommandation intégré à AI-VT utilisant un modèle Bayésien. Ce chapitre montre de quelle manière il permet de réviser une séance d'entraînement en cours. Le chapitre 8 montre l'utilisation du processus de Hawkes pour renforcer l'apprentissage.

Enfin, une conclusion générale incluant des perspectives de recherche termine ce manuscrit.

2

CONTEXTE

2.1/ NOTIONS ET ALGORITHMES

Dans ce chapitre sont décrits plus en détails les concepts et algorithmes utilisés dans le développement des modules. CEs modules font partie des contributions de cette thèse à l'environnement informatique pour l'apprentissage humain (EIAH) appelé AI-VT (*Artificial Intelligence - Artificial Trainer*).

2.1.1/ ENVIRONNEMENTS INFORMATIQUES POUR L'APPRENTISSAGE HUMAIN (EIAH)

Les EIAH sont des outils pour aider les apprenants dans le processus d'apprentissage et d'acquisition de la connaissance. Ces outils sont conçus pour fonctionner avec des stratégies d'apprentissage online et mixtes. Fondamentalement, l'architecture de ces systèmes est divisé en quatre composants comme le montre la figure 2.1 :

— Le domaine, qui contient les concepts, les règles et les stratégies pour résoudre des problèmes dans un domaine spécifique. Ce module peut détecter et corriger les erreurs des apprenants. L'information, généralement, est organisée selon des séquences pédagogiques.
— L'apprenant, ce module est le noyau du système car il contient toute l'information cognitive possible de l'apprenant ainsi que son évolution. Ce module doit avoir les informations implicites et explicites pour pouvoir créer une représentation de la connaissance et le processus d'apprentissage. C'est également important parce que cette information doit permettre de générer un diagnostic de l'état de la connaissance de l'apprenant, à partir duquel le système peut prédire les résultats dans certains domaines et choisir la stratégie optimale pour présenter les nouveaux contenus.
— L'enseignant, reçoit l'information des modules précédents grâce à laquelle il peut

prendre des décisions sur le changement de parcours ou sur la stratégie d'apprentissage. Il peut également interagir avec l'apprenant.

— L'interface, est le module qui gère les configurations et les interactions entre les modules par différents moyens interactifs.

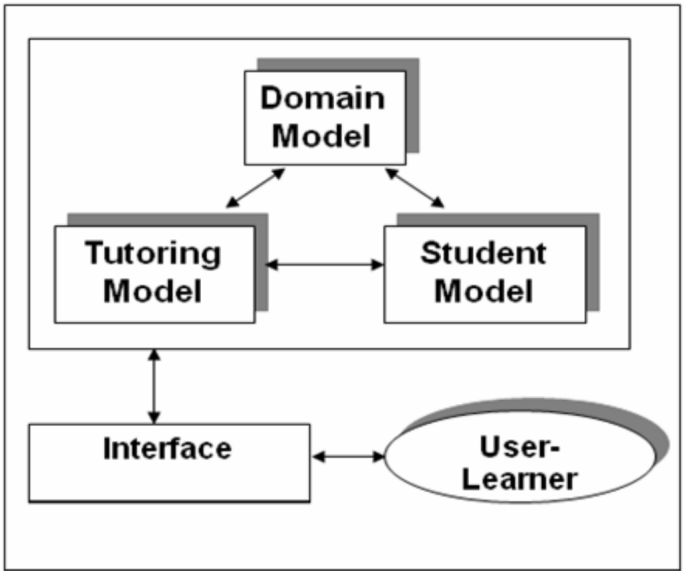

FIGURE 2.1 – L'architecture générale des EIAH, les composantes et leurs interactions ([Nkambou et al., 2010])

Le développement et la configuration de ce type de systèmes sont multidisciplinaires et impliquent plusieurs domaines de recherche comme : l'intelligence artificielle, la pédagogie, la psychologie, les sciences cognitives et l'éducation.

2.1.2/ STRATÉGIES D'APPRENTISSAGE

Dans [Lalitha and Sreeja, 2020], l'apprentissage est défini comme une fonction du cerveau humain acquise grâce au changement permanent dans l'obtention de connaissances et de compétences par le biais d'un processus de transformation du comportement lié à l'expérience ou à la pratique. L'apprentissage implique réflexion, compréhension, raisonnement et mise en oeuvre. Un individu peut utiliser différentes stratégies pour acquérir la connaissance. Comme le montre la figure 2.2, les stratégies peuvent être classées entre mode traditionnel et mode online.

L'apprentissage traditionnel se réfère au type d'apprentissage en classe, centré sur l'enseignant où la présence physique dans le temps et l'espace est fondamentale. Ici, l'enseignement est direct de l'enseignant à l'élève, et les ressources sont généralement

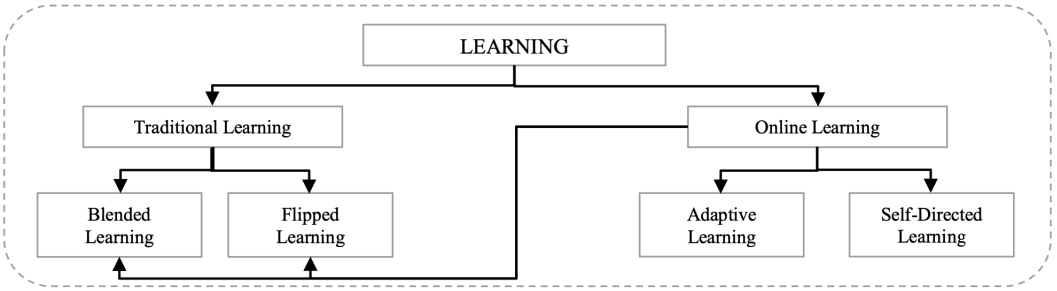

FIGURE 2.2 – Stratégies d'apprentissage ([Lalitha and Sreeja, 2020])

des documents imprimés. L'apprentissage et la participation doivent être actifs, la rétroaction est instantanée et des interactions sociales sont présentes. Par contre, il y a un manque de flexibilité en fonction du temps MLB :à revoir .

L'autre stratégie globale est l'apprentissage online, qui est basé sur les ressources et l'information disponible sur le web. Cette stratégie incite activement les apprenants à acquérir de nouvelles connaissances indépendamment MLB :indépendantes ou indépendemment à qq chose grâce aux vastes ressources disponibles en ligne. Les points positifs de cette stratégie sont par exemple : la rentabilité, la mise à niveau continue des compétences et des connaissances ou une plus grande opportunité d'accéder au contenu du monde entier. Pour l'apprenant, l'auto-motivation et l'interaction peuvent être des défis à surmonter. Pour l'enseignant, il est parfois difficile d'évaluer l'évolution du processus d'apprentissage de l'individu. L'apprentissage online fait aussi référence à l'utilisation de dispositifs électroniques pour apprendre (e-learning) où les apprenants peuvent être guidés par l'enseignant à travers des liens, du matériel spécifique d'apprentissage, des activités ou exercices.

Il existe aussi certaines stratégies mixtes qui combinent des caractéristiques des deux stratégies fondamentales, comme un cours online mais avec quelques séances en présentiel pour des activités spécifiques ou des cours traditionnels avec du matériel d'apprentissage complémentaire online.

2.1.3/ RAISONNEMENT À PARTIR DE CAS (RÀPC)

Le raisonnement à partir de cas est un type de modèle de raisonnement qui essaie d'imiter le comportement humain pour résoudre de nouveaux problèmes, en se souvenant des expériences passées. Le principe de base est que les problèmes similaires ont des solutions similaires. Si un nouveau problème arrive et que le modèle connait déjà la solution d'un problème similaire, alors

la solution au nouveau problème peut être inférée des solutions déjà trouvées MLB : suppr la fin de la phrase, redondant ?ou connues pour les problèmes similaires [Roldan Reyes et al., 2015].

Un cas est défini comme la représentation de la description d'un problème et la description de sa solution. Dans le raisonnement à partir de cas il existe une base de cas qui permet la création de nouvelles solutions en cherchant les problèmes similaires, en utilisant les données de la description du problème et leurs solutions associées.

Formellement, le modèle nécessite les ensembles P qui correspond à l'espace des problèmes et S l'espace des solutions, alors un problème x et sa solution y appartiennent à ces espaces $x \in P$ et $y \in S$. Si y est solution de x alors, un cas dans le modèle est représenté par le couple $(x, y) \in P \times S$. Le RàPC a besoin d'une base de n problèmes et de leurs solutions associées (x^n, y^n), cette base est appelée base de cas. L'objectif du RàPC est, étant donné un nouveau problème x^z de trouver sa solution y^z en utilisant la base de cas [Lepage et al., 2020].

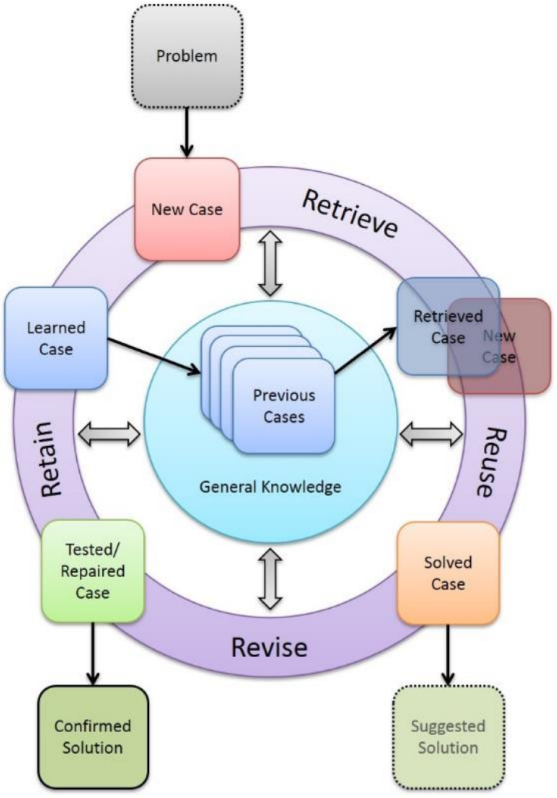

FIGURE 2.3 – Cycle fondamental du raisonnement à partir de cas ([Leikola et al., 2018])

Le processus du RàPC est divisé en quatre étapes qui forment un cycle, la figure 2.5

montre le cycle et le flux d'informations ainsi que les associations de chacune des étapes [Leikola et al., 2018] : rechercher, adapter, réviser et stocker. Chacune des étapes est décrite dans [Richter and Weber, 2013].

2.1.3.1/ RECHERCHER

L'objectif de cette étape est de rechercher dans la base de cas, les cas similaires à un nouveau cas donné. Cette similarité n'est pas un concept général, mais dépend du contexte, de l'objectif et du type de données. La question fondamentale qui se pose dans cette étape est : quel est le cas le plus approprié dans la base de cas qui me permet de réutiliser sa solution pour résoudre le nouveau problème donné ?

Pour évaluer le nouveau cas avec les cas de la base de données, il est nécessaire de les comparer et ainsi trouver la similitude entre eux. La comparaison entre les cas est différente selon la représentation du cas, c'est-à-dire la structure qui contient la description du problème. Généralement, il est MLB :Je ne comprends pas employé la représentation par attributs, qui décrit le cas (problème) comme un ensemble d'attributs avec des valeurs spécifiques, cette représentation est connue comme la représentation attribut-valeur. La similitude entre deux cas attribut-valeur $c_1 = a_{1,1}, a_{1,1}, ..., a_{1,n}$ et $c_2 = a_{2,1}, a_{2,1}, ..., a_{2,n}$ mesure la somme de la distance entre les attributs individuellement avec une valeur de pondération qui détermine l'importance de chaque attribut dans le mesure global comme le montre l'équation MLB :A revoir 2.1. Dans cette étape, il est possible aussi de sélectionner k différents cas similaires au nouveau cas donné.

$$\sum_{i=1}^{n} \omega_i \; sim(a_{1,i}, a_{2,i}) \tag{2.1}$$

2.1.3.2/ ADAPTER (RÉUTILISER)

L'idée du RàPC est d'utiliser l'expérience pour essayer de résoudre de nouveaux cas (problèmes). La figure 2.4 montre le principe de réutilisation où est utilisée la solution d'un cas (problème) ancien pour l'adapter et servir comme solution d'un nouveau cas.

Il existe des situations pour lesquelles le nouveau cas et un ancien cas sont très similaires alors la solution au cas ancien est aussi valide pour le nouveau cas. Cependant, le plus souvent il est nécessaire d'adapter l'ancienne solution pour arriver à une solution satisfaisante du nouveau cas. De la même manière que pour la phase 'rechercher', il existe beaucoup de stratégies d'adaptation qui dépendent de la représentation,

du contexte et des algorithmes utilisés. Une des stratégies les plus utilisées est la définition d'un ensemble de règles qui permettent la transformation de la solution MLB :ne comprends pas ou les solutions trouvées. L'objectif ici est d'inférer une nouvelle solution y_z du cas x_z à partir des k cas retrouvés dans la première étape.

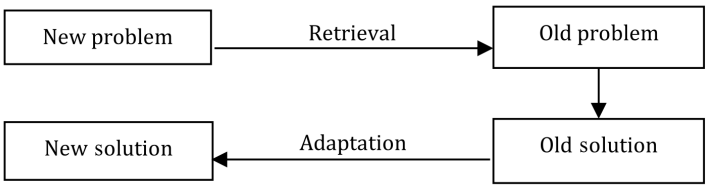

FIGURE 2.4 – Principe de réutilisation dans le RàPC ([Richter and Weber, 2013])

2.1.3.3/ RÉVISER ET RÉPARER

Une fois que le système propose une solution au nouveau cas, cette solution doit être évaluée et révisée qu'elle soit correcte selon certains critères de validation. L'étape de révision est chargée d'évaluer l'applicabilité de la solution obtenue dans l'étape 'adaptation', cette évaluation peut être réalisée dans le monde réel ou dans une simulation.

Alors dans cette phase, la tâche est de valider si le couple (x_z, y_z) est correct, c'est-à-dire si la solution trouvée y_z résout le problème x_z et satisfait les règles de validité et d'applicabilité. Si la solution n'est pas bonne en termes de qualité de la réponse, alors une nouvelle solution doit être générée en revenant à l'étape précédente. Si la solution est bonne, mais qu'elle ne satisfait pas les critères d'acceptation alors la solution doit être réparée pour corriger ses failles.

Le processus d'évaluation, de révision et de réparation entraine parfois un processus d'apprentissage qui permet d'améliorer la détection et la correction des failles des futures solutions générées. L'apprentissage est possible, car les modèles d'évaluation/réparation utilisent souvent des paramètres pour donner des résultats et sur ces paramètres il est possible d'appliquer des algorithmes qui vont aider à mesurer et changer les solutions selon les cas particuliers.

L'évaluation peut se faire soit par un expert humain qui a la capacité d'évaluer la solution selon le contexte et la pertinence, soit par un système qui fonctionne dans le monde réel renvoyant le résultat de l'application de la solution ou soit par un modèle qui utilise des simulations numériques ou données et des tests statistiques. La correction des

solutions peut être automatique, mais dépend du domaine et de l'information disponible pour le faire, usuellement le conteneur du vocabulaire a un ensemble de règles qui sont appliquées aux solutions qui ont été identifiées comme non valides.

2.1.3.4/ STOCKER (RETENIR)

Si un nouveau cas résolu est jugé pertinent, alors celui-ci peut être stocké dans la base de cas pour qu'il aide dans la résolution de futures nouveaux cas. Dans le RàPC c'est considéré comme un processus d'apprentissage, car tous les cas résolus ne sont pas retenus, seulement certains d'entre eux, ceux qui remplissent des critères prédéfinis. Cette étape a deux rôles différents, le premier est d'aider le système à améliorer sa capacité à sélectionner des cas, évaluer et corriger les solutions générées ; et le second est d'optimiser la taille de la base de cas pour éviter de gaspiller l'espace de stockage et ralentir le système au moment de trouver les cas similaires.

2.1.3.5/ CONTENEURS DE CONNAISSANCE

Pour pouvoir exécuter le cycle complet, le RàPC dépend de quatre sources différentes d'information, ces sources sont appelées les conteneurs. Les conteneurs sont : le conteneur de cas, le conteneur d'adaptation, le conteneur du vocabulaire et le conteneur de similarité [Richter, 2009] :

— Conteneur de cas : il contient les expériences passées que le système peut utiliser pour résoudre les nouveaux problèmes. L'information est structurée comme un couple (p, s), où p est la description d'un problème et s est la description de sa solution.
— Conteneur d'adaptation : il stocke le ou les stratégies d'adaptation ainsi que les règles et paramètres nécessaires pour les exécuter.
— Conteneur du vocabulaire : il contient l'information de la représentation. MLB :c'est-à-dire ?les éléments... les fonctions ? Les éléments comme les entités, les attributs, les fonctionnes ou les relations entre les entités. Généralement dans le RàPC est utilisée la représentation des cas avec la structure clé-valeur.
— Conteneur de similarité : il stoke les paramètres et la mesure de similarité $sim(p_1, p_2)$ entre les cas p_1 et p_2. MLB :ne comprends pas la phrase De plus, il est nécessaire de définir la fonction comme si le résultat est grande, alors la similitude entre p_1 et p_2 est importante ou au contraire.

La figure 2.5 montre les liens entre les étapes du RàPC et les conteneurs, les flèches

montrent le flux de l'information, les lignes continues suivent le parcours du cycle du RàPC, les lignes discontinues sont les liens entre les étapes et les conteneurs. L'étape "Retain" peut elle interagir avec les conteneurs 2, 3 et 4 comme dans le travail de Marie [Marie, 2019].

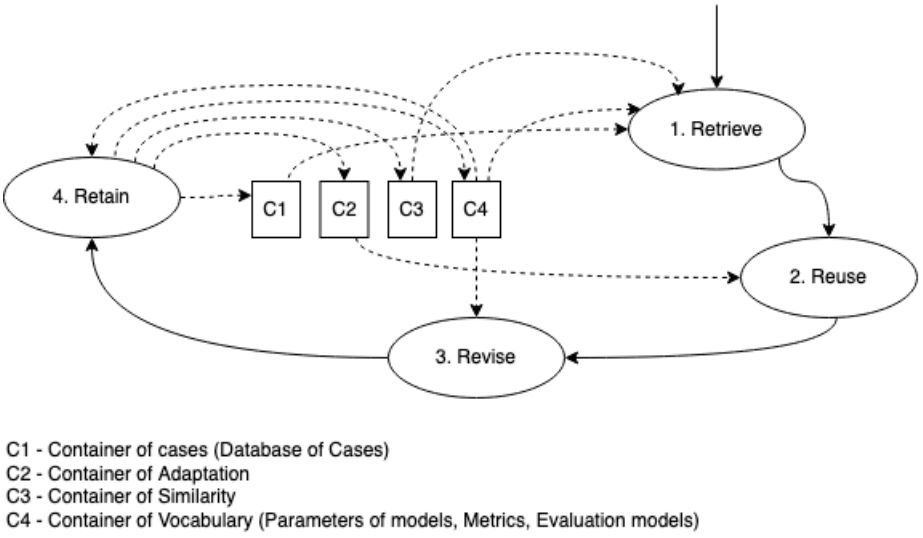

FIGURE 2.5 – Cycle du RàPC, les étapes, les conteneurs et leurs liens

2.1.4/ SYSTÈMES MULTI-AGENTS

Les systèmes multi-agents sont des systèmes conçus pour résoudre des problèmes en combinant l'intelligence artificielle et le calcul distribué. Ces systèmes sont composés de multiples entités autonomes appelées agents, qui ont la capacité de communiquer entre elles et également de coordonner leurs comportements [Hajduk et al., 2019].

Les agents ont comme propriétés :

- Autonomie : les agents fonctionnent sans intervention externe des êtres humains ou d'autres entités, ils ont un mécanisme interne qui leur permet de contrôler leurs états.
- Réactivité : les agents ont la capacité de percevoir l'environnement dans lequel ils sont et peuvent réagir aux changements qui se produisent.
- Pro-activité : les changements et réactions peuvent se produire aussi comme l'effet du processus cognitif interne, non seulement à cause des changements dans l'environnement MLB :phrase à revoir .
- Coopération : étant donné que les agents peuvent communiquer les uns avec les autres, ils échangent de l'information qui les aide à se coordonner et à résoudre

un même problème.
— Apprentissage : il est nécessaire qu'un agent soit capable de réagir dans un environnement dynamique et inconnu, il doit donc avoir la capacité d'apprendre de ses interactions et ainsi améliorer la qualité de ses réactions et comportements.

Il existe quatre types d'agent en fonction des capacités et des approches :
— Réactif : c'est l'agent qui perçoit constamment l'environnement et agit en fonction de ses objectifs.
— Basé sur les réflexes : c'est l'agent qui considère les options pour atteindre ses objectifs et développe un plan à suivre.
— Hybride agent : il combine les deux modèles antérieurs en utilisant chacun d'eux en fonction de la situation et de l'objectif.
— Basé sur le comportement : l'agent contient un ensemble de modèles de comportement pour réaliser certaines tâches spécifiques, chaque comportement se déclenche selon des règles prédéfinies ou des conditions d'activation. Le comportement de l'agent peut être modélisé avec différentes stratégies cognitives de pensée ou de raisonnement.

2.1.5/ PENSÉE BAYESIENNE

D'après les travaux de certains mathématiciens et neuroscientifiques, toute forme de cognition peut être modélisée dans le cadre de la formule de Bayes (equation 2.2), car elle permet de tester différentes hypothèses et donner plus de crédibilité à celle qui est confirmée par les observations. Étant donné que ce type de modèle cognitif permet l'apprentissage optimal, la prédiction sur les événements passés, l'échantillonnage représentatif et l'inférence d'information manquante ; il est appliqué dans quelques algorithmes de machine learning et d'intelligence artificielle [Hoang, 2018].

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2.2}$$

La formule de Bayes calcule la probabilité a posteriori $P(A|B)$ de la plausibilité de la théorie A compte tenu des données B, et requiert trois termes : (1) une probabilité a priori $P(A)$ de la plausibilité de l'hypothèse A, (2) le terme $P(B|A)$ qui mesure la capacité de l'hypothèse A à expliquer les données observées B et (3) la fonction de partition $P(B)$ qui met en compétition toutes les hypothèses qui ont pu générer les données observées B. À chaque nouvelle évaluation de la formule, la valeur du terme a priori $P(A)$ est actualisée par la valeur du terme a posteriori $P(A|B)$. Ainsi, à chaque évaluation, le degré de plausibilité de chaque hypothèse est ajusté [Hoang, 2018].

Par la suite, nous explicitons quelques algorithmes, intègrés généralement dans l'étape "Rechercher" du RàPC, pour la recherche des cas les plus proches d'un nouveau cas.

2.1.6/ MÉTHODE DES K PLUS PROCHES VOISINS (K-NEAREST NEIGHBORHOOD - KNN)

Comme est défini dans [Cunningham and Delany, 2021], la méthode des 'k' plus proches voisins est une méthode pour classifier des données dans des classes spécifiques. Pour faire cela, il est nécessaire de disposer d'une base de données avec des exemples déjà identifiés. Pour classer de nouveaux exemples, la méthode attribue la même classe que celle de leurs voisins les plus proches. Généralement, l'algorithme compare les caractéristiques d'une entité avec plusieurs possibles voisins pour essayer d'btenir des résultats plus précis. 'k' représente le nombre de voisins, puisque l'algorithme peut être exécuté à chaque fois avec un nombre différent de voisins.

Étant donné un jeu de données D constitue de $(x_i)_{i\in[1,n]}$ données (où $n = |D|$). Chacune des données est décrite par F caractéristiques qui sont des valeurs numériques normalisées $[0, 1]$, et par une classe de labellisation $y_j \in Y$. Le but est de classifier une donnée inconnue q. Pour chaque $x_i \in D$, il est possible de calculer la distance entre q et x_i comme l'équation 2.3.

$$d(q, x_i) = \sum_{f\in F} w_f \delta(q_f, x_{if}) \tag{2.3}$$

C'est une somme de MLB :à vérifier de toutes les caractéristiques F avec w_f le poids pour chaque caractéristique. MLB :vérifier la suite La fonction de distance δ peut être une métrique générique. Pour des valeurs numériques discrètes il est possible d'utiliser la définition 2.4,

$$\delta(q_f, x_{if}) = \begin{cases} 0 & q_f = x_{if} \\ 1 & q_f \neq x_{if} \end{cases} \tag{2.4}$$

et si les valeurs sont continues l'équation 2.5.

$$\delta(q_f, x_{if}) = |q_f - x_{if}| \tag{2.5}$$

MLB :paragraphe A vérifier, dernière phrase incomprise Souvent, pour décider de la classe d'une donnée inconnue, on attribue un poids plus important aux voisins les plus proches. Une technique assez générale pour y parvenir est le vote pondéré en fonction de

la distance. Les voisins peuvent voter sur une certaine classe pour attribuer à la donnée inconnue avec des votes pondérés par l'inverse de leur distance (équation 2.6).

$$N(a,b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \tag{2.6}$$

L'équation de vote peut être comme l'équation 2.7.

$$vote(y_i) = \sum_{c=1}^{k} \frac{1}{d(q,x_c)^p} N(y_j, y_c) \tag{2.7}$$

une autre alternative est basée sur le travail de Shepard, équation 2.8.

$$vote(y_i) = \sum_{c=1}^{k} e^{d(q,x_c)} N(y_j, y_c) \tag{2.8}$$

La fonction de distance peut être n'importe quelle mesure d'affinité entre deux objets, mais cette fonction doit répondre à quatre critères (équations 2.9).

$$\begin{aligned} &d(x,y) \geq 0 \\ &d(x,y) = 0, \ \ seulement\ si\ x = y \\ &d(x,y) = d(y,x) \\ &d(x,z) \geq d(x,y) + d(y,z) \end{aligned} \tag{2.9}$$

2.1.7/ K-MOYENNES

Selon [Sinaga and Yang, 2020], K-Moyennes est un algorithme d'apprentissage utilisé pour partitionner et grouper des données. MLB :il faur un verbe dans la phrase Étant donné $X = \{x_1, ..., x_n\}$ un ensemble dans un espace Euclidien $\mathbb{R}^d$. Et $A = \{a_1, ..., a_c\}$ avec c comme nombre de groupes. Ainsi que $z = [z_{ik}]_{n \times c}$, où z_{ik} est une variable binaire (i.e. $z_{ik} \in \{0, 1\}$) qui indique si une donnée x_i appartient au k-ème groupe, $k = 1, , c$. L'équation 2.10 montre la fonction objective k-moyenne.

$$J(z,A) = \sum_{i=1}^{n} \sum_{k=1}^{c} z_{ik} \parallel x_i - a_k \parallel^2 \tag{2.10}$$

MLB :a discuter L'algorithme des k moyennes est itéré à travers les conditions nécessaires pour minimiser la fonction objectif des k moyennes $J(z,A)$ avec la mise à jour des équations pour les centres du groupe (équation 2.11) et les appartenances (2.12), respectivement.

$$a_k = \frac{\sum_{i=1}^{n} z_{ik} x_{ij}}{\sum_{i=1}^{n} z_{ik}} \tag{2.11}$$

$$z_{ik} = \begin{cases} 1 & if \parallel x_i - a_k \parallel^2 = min_{1 \leq k \leq c} \parallel x_i - a_k \parallel^2 \\ 0, & otherwise \end{cases} \tag{2.12}$$

2.1.8/ MODÈLE DE MÉLANGE GAUSSIEN GMM (*Gaussian Mixture Model*)

Comme décrit dans [Wang et al., 2021], le modèle de mélange gaussien (GMM) est un modèle probabiliste composé de plusieurs modèles gaussiens simples. Considérant une variable multidimensionnelle $x = \{x_1, x_2, ..., x_d\}$ en tant qu'entrée, GMM multivarié est défini comme dans l'équation 2.13.

$$p(x|\theta) = \sum_{k=1}^{K} \alpha_k g(x|\theta_k) \tag{2.13}$$

Dans cette équation, K est le nombre de modèles gaussiens uniques dans GMM, également appelés composants ; MLB :pas compris α_k est la probabilité de mélange de la k-ème composante du modèle de mélange sujette à la condition que montre l'équation 2.14.

$$\sum_{k=1}^{K} \alpha_k = 1 \tag{2.14}$$

θ_k représente la valeur moyenne et la matrice de covariance de chaque modèle gaussien unique : $\theta_k = \{\mu_k, \Sigma_k\}$. Pour un seul modèle gaussien multivarié, nous avons la fonction de densité de probabilité $g(x)$ (équation 2.15).

$$g(x; \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right) \tag{2.15}$$

La tâche principale est d'obtenir les paramètres α_k, θ_k pour tout k défini dans GMM en utilisant un ensemble de données avec N échantillons d'entrainement. Une solution classique et disponible pour l'estimation des paramètres requis utilise l'algorithme de maximisation des attentes (Expectation-Maximization EM), qui vise à maximiser la vraisemblance de l'ensemble de données. Il s'agit d'une approche itérative, et les paramètres sont continuellement mis à jour jusqu'à ce qu'ils répondent au critère terminal : MLB :la suite est bien l'expression du critère ? lorsque la valeur delta log-vraisemblance entre deux itérations est inférieure à un seuil donné. Les paramètres du modèle de mélange peuvent enfin être obtenus après des itérations répétées, qui présentent une

estimation de la distribution de l'ensemble des données et qui MLB :ok qui ? sont utilisés pour une opération de génération ultérieure.

2.1.9/ FUZZY-C

Fuzzy C-Means Clustering (FCM) est un algorithme de clustering flou non supervisé largement utilisé [Xu et al., 2021]. Le FCM utilise comme mesure de distance la mesure euclidienne. Supposons d'abord que l'ensemble d'échantillons à regrouper est $X = \{x_1, x_2, ..., x_n\}$, où $x_j \in R^d (1 \leq j \leq n)$ dans le d-dimensionnel espace Euclidien, et c est le nombre de clusters. L'équation 2.16 montre la fonction objectif de FCM.

$$J(U, V) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} (x_j - v_i)^T A (x_j - v_i) \tag{2.16}$$

où A est la matrice métrique, m est un nombre quelconque ($m > 1$) qui dénote le degré de flou, u_{ij} est le degré d'appartenance du j-ème échantillon x_j qui appartient au i-ème cluster, dont le centre est v_i. $U = (u_{ij})$, $V = [v_1, v_2, ..., v_c], 1 \leq i \leq c, 1 \leq j \leq n, 2 \leq c < n$, et u_{ij} satisfait les conditions de l'équation 2.17.

$$\sum_{i=1}^{c} u_{ij} = 1, u_{ij} \geq 0 \tag{2.17}$$

Pour compléter le modèle, les équations de mise à jour pour le centre du cluster v_i (équation 2.18) et des degrés d'appartenance u_{ij} (équation 2.19) sont définies.

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}} \tag{2.18}$$

$$u_{ij} = \frac{((x_j - v_i)^T A (x_j - v_i))^{\frac{2}{m-1}}}{\left(\sum_{h=1}^{c} (x_j - v_h)^T A (x_j - v_h)\right)^{\frac{2}{m-1}}} \tag{2.19}$$

Les algorithmes qui se trouvent dans la suite de cette section sont des algorithmes qui font partie de l'intelligence artificielle et sont utilisés dans certains EIAH pour améliorer les recommandations, essayer de corriger et de détecter les faiblesses des apprenants de façon automatique.

2.1.10/ BANDIT MANCHOT MAB (*Multi-Armed Bandits*)

Dans [Gupta et al., 2021] MAB est décrit comme un problème qui appartient au domaine de l'apprentissage par renforcement, celui-ci représente un processus de prise

de décision séquentielle dans des instants t de temps, où un utilisateur doit sélectionner une action $k_t \in K$ à réaliser dans un ensemble K fini de possibles actions et chacune d'elles donne une récompense qui est inconnue pour l'utilisateur. L'objectif est de maximiser la récompense cumulée globale dans le temps. La clé pour trouver une solution au problème est de trouver l'équilibre optimal entre l'exploration (exécuter des actions inconnues pour collecter de l'information complémentaire) et l'exploitation (continuer à exécuter les actions déjà connues pour cumuler plus de gains). L'un des algorithmes utilisés pour résoudre ce problème c'est l'échantillonnage de Thompson.

2.1.11/ ÉCHANTILLONNAGE DE THOMPSON TS (*Thompson Sampling*)

Comme indiqué dans [Lin, 2022], l'algorithme d'échantillonnage de Thompson est un algorithme de type probabiliste utilisé généralement pour résoudre le problème MAB. MLB :ne comprends pas la suite Il est basé sur le principe Bayesien où il y a une distribution de probabilité a priori et avec les données obtenues est génèrée une distribution de probabilité a posteriori utilisée pour essayer de maximiser l'estimation de la valeur espérée, la distribution de base utilisée est la distribution Beta qui est définie sur $[0, 1]$ et paramétrée par deux valeurs α et β, c'est un cas particulier de la loi de Dirichlet comme le montre la figure 2.6. L'équation 2.20 décrit formellement la famille des courbes générées par la distribution Beta.

Chaque action de MAB a une distribution Beta dont les paramètres sont initialisés en 1. Ces valeurs changent et sont calculées à partir des récompenses obtenues : si au moment d'exécuter une action spécifique le résultat est un succès, alors la valeur du paramètre α de sa distribution Beta doit augmenter mais si le résultat est un échec alors la valeur du paramètre β de sa distribution Beta doit augmenter. De ce façon, la distribution pour chacune des possibles actions est ajustée en privilégiant les actions qui génèrent le plus de récompenses sur les autres actions. A chaque nouvelle itération, toutes les distributions sont plus précises sur la valeur espérée des actions.

$$B(x, \alpha, \beta) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du} & pour x \in [0, 1] \\ 0 & sinon \end{cases} \tag{2.20}$$

Dans les deux chapitres suivants nous présentons les travaux récents, en rapport avec le sujet de thèse, qui utilisent les concepts et algorithmes explicités dans ce chapitre.

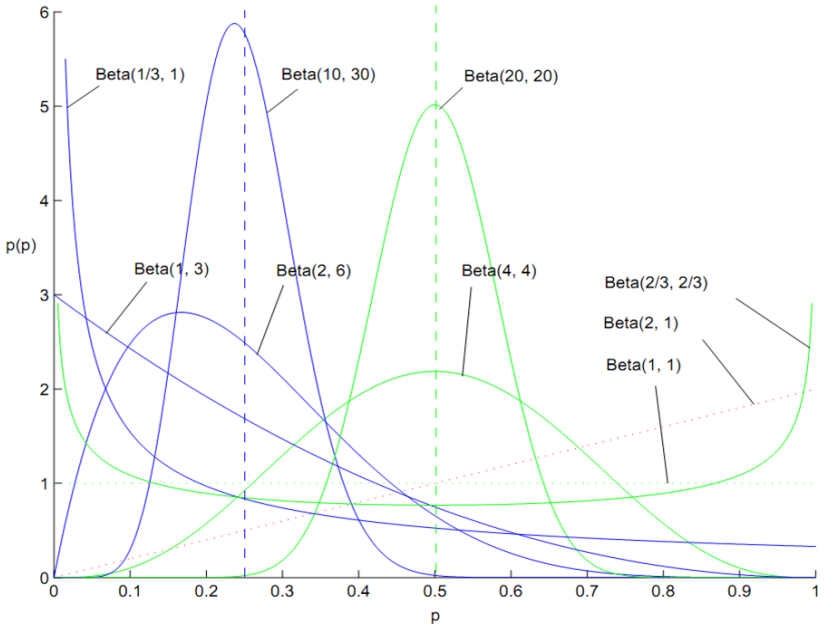

FIGURE 2.6 – Comportement de la distribution Beta avec différents valeurs dans ses paramètres

II

ÉTAT DE L'ART

3

Environnements Informatiques d'Apprentissage Humain

Dans ce chapitre sont mentionnés des travaux qui sont en rapport avec le travail de la thèse spécifiquement sur les EIAH, mais ils utilisent des approches différents, alors ici est faite une classification de ces travaux mais selon le thème principal qu'ils abordent.

3.1/ L'Intelligence Artificielle

L'intelligence artificielle a été appliquée avec succès au domaine de l'éducation dans plusieurs écoles de différents pays et pour l'apprentissage de l'informatique, les langues étrangères et les sciences comme le montre le travail de [Zhang. and Aslan, 2021] qui analyse des articles dans la période de 1993-2020, où il est importante de voir aussi que il n'y a pas une distinction en age, ni en niveau culturel, ni en niveau éducatif. En plus, avec les différentes techniques d'intelligence artificielle qui sont appliquées il est possible d'adapter la stratégie d'apprentissage pour n'importe qui et maximiser le rendement et acquisition des connaissances. Une caractéristique assez notable est que les représentations et algorithmes peuvent changer et évoluer. Aujourd'hui il y a encore beaucoup de potentiel pour développer encore les capacités de calcul et d'adaptation.

Un travail plus analytique de l'utilisation de l'intelligence artificielle dans l'éducation est [Chiu et al., 2023] qui présente une révision des travaux d'intelligence artificielle appliquée à l'éducation extraits des bases de données bibliographiques ERIC, WOS et Scopus dont la date de publication est dans la période 2012-2021, ce travail est focalisé sur les tendances et les outils employées pour aider dans le processus d'apprentissage. Après l'analyse de 92 travaux, une classification des contributions de l'IA est crée : apprentissage, enseignement, évaluation et administration. Dans l'apprentissage, est utilisée généralement l'IA pour attribuer des tâches en fonction des

compétences individuelles, fournir des conversations homme-machine, analyser le travail des étudiants pour obtenir des commentaires et accroître l'adaptabilité et l'interactivité dans les environnements numériques. Dans l'enseignement, peuvent être appliquées des stratégies d'enseignement adaptatives, améliorer la capacité des enseignants à enseigner et soutenir le développement professionnel des enseignants. Dans le domaine de l'évaluation, l'IA peut fournir une notation automatique et prédire les performances des élèves. L'IA aide dans le domaine de l'administration à améliorer la performance des plateformes de gestion, à soutenir la prise de décision pédagogique et à fournir des services personnalisés. Quelques-unes des lacunes identifiées de l'IA dans le domaine de l'éducation sont : les ressources recommandées par les plateformes d'apprentissage personnalisées sont trop homogènes, les données nécessaires pour les modèles d'IA sont très spécifiques, le lien entre les technologies et l'enseignement n'est pas bien clair, la plupart des travaux sont conçus pour un domaine ou un objectif très spécifique (peu de généralité), l'inégalité éducative car les technologies ne motivent pas tous les types d'élèves et parfois il y a des attitudes négatives envers l'IA, parce qu'elle est difficile à maîtriser et a intégrer dans les cours.

Les techniques d'IA peuvent aussi aider à prendre des décisions stratégiques qui visent des objectifs à long échéance comme le montre le travail de [Robertson and Watson, 2014] qui analyse plusieurs publications réalisées dans le domaine de l'utilisation de l'IA en jeux stratégiques. Les jeux stratégiques sont un très bon terrain d'entrainement parce que ils contient plusieurs règles, environnements pleins de contraintes, actions et réactions en temps réel, caractère aléatoire et informations cachées. Les techniques principales identifiées sont : l'apprentissage par renforcement, les modèles Bayésiens, la recherche en arbre, le raisonnement à partir de cas, les réseaux de neurones et les algorithmes évolutifs. Ici est notable aussi comme la combinaison de ces techniques permet dans certains cas d'améliorer le comportement global des algorithmes et d'obtenir meilleures réponses.

3.2/ Systèmes de Recommandation dans les EIAH

Les systèmes de recommandation dans les environnements d'apprentissage considèrent les exigences, les nécessites, le profil, les talents, les intérêts et l'évolution de l'apprenant pour s'adapter et recommander des ressources ou des exercices avec le but d'améliorer l'acquisition et maîtrise de concepts et des connaissances en général. L'adaptation de ces systèmes peut être de deux types, l'adaptation de la présentation qui montre aux apprenants les ressources en concordance avec leurs faiblesses et l'adaptation de la navigation qui change la structure du cours en fonction du niveau et style d'apprentissage

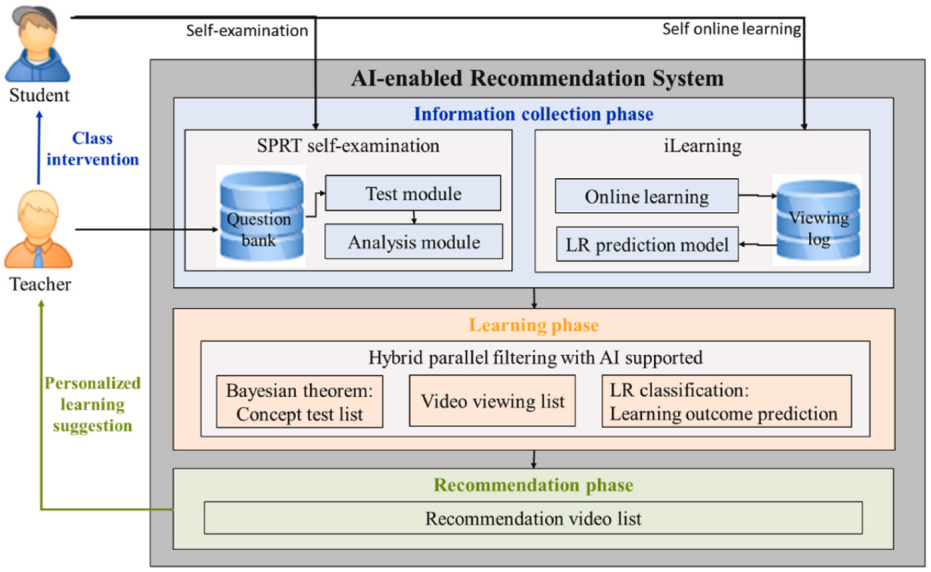

FIGURE 3.1 – Architecture du système de recommandation proposé dans [Huang et al., 2023]

de chaque apprenant [Muangprathub et al., 2020]. Ces systèmes montrent des effets positifs pour les apprenants comme le révèle le travail de [Huang et al., 2023], qui applique l'intelligence artificielle pour personnaliser des recommandations de ressources en vidéo et ainsi aider et motiver les apprenants, les effets ont été validés avec la différence entre des test préliminaires et des test après la finalisation du cours ainsi que un groupe de contrôle qui a suivi le même cours sans le système de recommandation. La figure 3.1 montre l'architecture du système où l'intelligence artificielle est utilisée dans différents étapes du processus avec les données de l'apprenant et la supervision et contrôle du professeur.

Le travail de [Seznec et al., 2020] propose un modèle générique de recommandation qui peut être utilisé dans les systèmes de recommandation de produits ou les systèmes d'apprentissage. Ce modèle est basé sur l'algorithme par renforcement UCB (Upper Confidence Bound) qui permet de trouver une solution approchée à une variante du problème bandit manchot non stationnaire (MAB), où les recompenses de chaque action diminuent chaque fois que elle est utilisée. Pour valider le modèle une comparaison avec trois autres algorithmes sur une base de données réelle a été réalisée, les métriques employées sont la moyenne accumulée du régret et la moyenne de la récompense accumulée. Avec les deux métriques, le modèle proposé est très performant.

Aussi [Ingkavara et al., 2022], met en évidence que les technologies et les systèmes de recommandation peuvent s'adapter à différents besoins et aspirations ainsi que

favoriser l'apprentissage auto-régulé. Ce type d'apprentissage aide aux apprenants à acquérir les habilités pour améliorer leur vitesse et performance ; parce que y trouvent des objectifs variables, un environnement structuré, des temps d'apprentissage variable et des ressources de support et de renforcement.

Comme vu précédemment, les techniques de l'intelligence artificielle sont largement utilisées dans l'apprentissage et l'enseignement ; l'apprentissage adaptatif pour suggérer des ressources d'étude est l'objectif du travail de [Lalitha and Sreeja, 2020] où est proposé un système avec un module de personnalisation qui collecte l'information de l'apprenant et ses exigences, un module de classification qui compare l'information avec d'autres profiles en utilisant l'algorithme KNN et un module de recommandation chargé de identifier les ressources communs entre les profiles et d'extraire information complémentaire d'Internet avec Random Forest pour améliorer le processus d'apprentissage du nouveau apprenant. Les techniques et les stratégies sont très variées mais l'objectif est de s'adapter aux apprenants et leurs besoins ; Dans [Zhao et al., 2023] les données des apprenants sont collectés et classifiés dans groupes avec des caractéristiques similaires, puis pour déterminer la performance de chaque apprenant est utilisée la méthode d'analyse par enveloppement des données (DEA) qui permet d'identifier les besoins spécifiques de chaque groupe et ainsi proposer un parcours d'apprentissage personnalisé.

Un aspect pertinent des systèmes de recommandation qu'il faut bien considérer est la représentation des données des enseignants, des apprenants et l'environnement ; parce que les multiples types de structures qu'on peut utiliser et son contenu peuvent conduire vers des résultats différents et changements dans la performance des algorithmes ou techniques employées. La proposition de [Su et al., 2022] est de stocker les données des apprenants dans deux graphes évolutifs qui contient les relations entre les questions et les réponses correctes ainsi que les réponses données par les apprenants, pour construire un graphe global de chaque apprenant avec son état cognitif et avec lequel est analysé et prédit la performance dans un contexte spécifique et ainsi pouvoir proposer un parcours personnalisé. Dans [Muangprathub et al., 2020] sont représentés les concepts comme trois ensembles : les objets (G), les attributs (M) et les relations entre G et M ; avec l'idée de les analyser avec l'approche FCA (Formal Context Analysis), Ce type de représentation permet de mettre en rapport les ressources, les questions et les sujets, ainsi si un apprenant étudie un sujet S_1 et il y a une règle qui relie S_1 avec le sujet S_4 ou la question $Q_3 4$, alors le système peut suggérer l'étude de ces ressources, l'algorithme complet explore la structure prédéfinie du cours pour recommander un parcours exhaustif d'apprentissage.

La recommandation personnalisé d'exercices est aussi une autre approche des systèmes de recommandation comme le fait [Zhou and Wang, 2021] de forme spécifique pour l'apprentissage de l'anglais, le système contient un module principal qui représente les apprenants comme des vecteurs selon le modèle DINA où sont stockes les points de connaissance acquise, le vecteur est n-dimensionnel $K = \{k_1, k_2, ..., k_n\}$ et chaque dimension correspond à un point de connaissance, ainsi si $k_1 = 1$ alors l'apprenant maîtrise le point de connaissance k_1, et si $k_2 = 0$ alors l'apprenant doit étudier le point de connaissance k_2 car il n'a pas acquise la maîtrise dans ce point-là. Sur ce type de représentation est basé la recommandation des questions proposées par le système, il y a aussi une librairie de ressources associes à chacune des possibles questions, avec lesquelles l'apprenant peut renforcer son apprentissage et avancer dans le programme du cours.

Certains travaux de recommandation et personnalisation considèrent des variables complémentaires aux notes comme dans [Ezaldeen et al., 2022], qui lie l'analyse du comportement de l'apprenant et l'analyse sémantique. La première étape consiste en collecter les données nécessaires pour créer un profile de l'apprenant, avec le profile complet, sont associés l'apprenant et un groupe de catégories prédéfinies d'apprentissage selon ses préférences et les données historiques, puis en ayant une valeur pour chaque catégorie sont cherchés les concepts associés pour les catégories et ainsi est générée une guide pour obtenir des ressources qu'on peut recommander sur le web. Le système est divisé en quatre niveaux comme montre la figure 3.2 où chacun des niveaux est chargé de une tache spécifique dont les résultats sont envoyés au niveau suivant jusqu'à arriver aux recommandations personnalisés pour l'apprenant.

Une limitation générale qui presentent les algorithmes d'IA dans les EIAH est que ce type d'algorithmes ne permettent pas d'oublier l'information avec laquelle ils ont été entrainés, une propiété nécessaire pour les EIAH dans certaines situations où peut se produire un disfonctionnement du système, la réévaluation d'un apprenant ou la reestructuration d'un cours.

Le tableau 3.1 montre un récapitulatif des articles analysés dans l'état de l'art des EIAH.

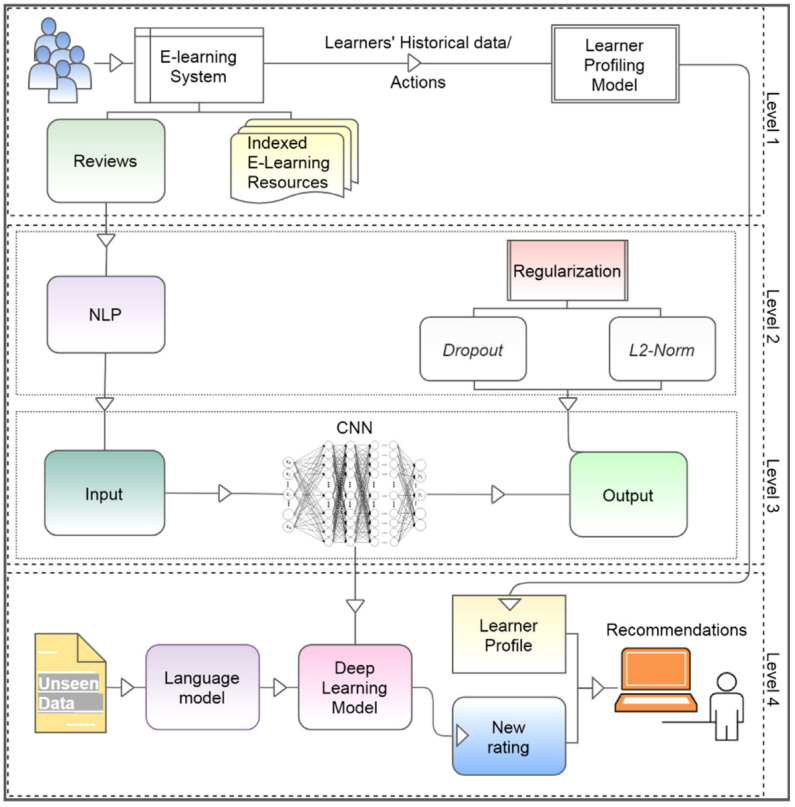

FIGURE 3.2 – Niveaux du système de recommandation dans [Ezaldeen et al., 2022]

Référence	Limites et faiblesses
[Zhang. and Aslan, 2021]	Niveau global d'application de EIAH. Analyse des effets de l'IA
[Chiu et al., 2023]	Aspects non cognitifs en IA. Motivation des apprenants
[Robertson and Watson, 2014]	Peu de test. Pas de standard d'évaluation
[Huang et al., 2023]	Recommandation en fonction de la motivation seulement. N'est pas général pour tous les apprenants
[Seznec et al., 2020]	Le modèle n'a pas été testé avec données d'étudiants. UCB prends beaucoup de temps pour explorer les alternatives
[Ingkavara et al., 2022]	Modèle très complexe. Définition de constantes subjectif et beaucoup de variables.
[Lalitha and Sreeja, 2020]	A besoin de beaucoup de données pour entrainement. Ne marche pas dans le cas 'cold-start', nést pas totalement automatisé
[Su et al., 2022]	Estimation de la connaissance de façon subjective. Il est nécessaire une étape d'entrenaiment
[Muangprathub et al., 2020]	Structure des règles complexe. Définition de la connaissance de base complexe
[Zhou and Wang, 2021]	Utilisation d'un filtre collaboratif sans stratification. Utilisation d'une seule metrique de distance
[Ezaldeen et al., 2022]	Il n y a pas de comparaison avec d'autres modèles différents de CNN. Beaucoup de variables à valeurs subjectives

TABLE 3.1 – Tableau de synthèse des articles analysés dans l'état de l'art des EIAH

BIBLIOGRAPHIE

[Chiu et al., 2023] Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., and Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education : Artificial Intelligence*, 4 :100118.

[Cunningham and Delany, 2021] Cunningham, P. and Delany, S. J. (2021). K-nearest neighbour classifiers - a tutorial. *ACM Comput. Surv.*, 54(6).

[Ezaldeen et al., 2022] Ezaldeen, H., Misra, R., Bisoy, S. K., Alatrash, R., and Priyadarshini, R. (2022). A hybrid e-learning recommendation integrating adaptive profiling and sentiment analysis. *Journal of Web Semantics*, 72 :100700.

[Gupta et al., 2021] Gupta, S., Chaudhari, S., Joshi, G., and Yağan, O. (2021). Multi-armed bandits with correlated arms. *IEEE Transactions on Information Theory*, 67(10) :6711–6732.

[Hajduk et al., 2019] Hajduk, M., Sukop, M., and Haun, M. (2019). *Cognitive Multi-agent Systems : Structures, Strategies and Applications to Mobile Robotics and Robosoccer*. Studies in Systems, Decision and Control. Springer International Publishing.

[Henriet et al., 2017] Henriet, J., Christophe, L., and Laurent, P. (2017). Artificial intelligence-virtual trainer : An educative system based on artificial intelligence and designed to produce varied and consistent training lessons. *Proceedings of the Institution of Mechanical Engineers, Part P : Journal of Sports Engineering and Technology*, 231(2) :110–124.

[Hoang, 2018] Hoang, L. (2018). *La formule du savoir. Une philosophie unifiée du savoir fondée sur le théorème de Bayes*. EDP Sciences.

[Huang et al., 2023] Huang, A. Y., Lu, O. H., and Yang, S. J. (2023). Effects of artificial intelligence–enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers and Education*, 194 :104684.

[Ingkavara et al., 2022] Ingkavara, T., Panjaburee, P., Srisawasdi, N., and Sajjapanroj, S. (2022). The use of a personalized learning approach to implementing self-regulated online learning. *Computers and Education : Artificial Intelligence*, 3 :100086.

[Lalitha and Sreeja, 2020] Lalitha, T. B. and Sreeja, P. S. (2020). Personalised self-directed learning recommendation system. *Procedia Computer Science*, 171 :583–

592. Third International Conference on Computing and Network Communications (CoCoNet'19).

[Leikola et al., 2018] Leikola, M., Sauer, C., Rintala, L., Aromaa, J., and Lundström, M. (2018). Assessing the similarity of cyanide-free gold leaching processes : A case-based reasoning application. *Minerals*, 8(10).

[Lepage et al., 2020] Lepage, Y., Lieber, J., Mornard, I., Nauer, E., Romary, J., and Sies, R. (2020). The french correction : When retrieval is harder to specify than adaptation. In Watson, I. and Weber, R., editors, *Case-Based Reasoning Research and Development*, pages 309–324, Cham. Springer International Publishing.

[Lin, 2022] Lin, B. (2022). Evolutionary multi-armed bandits with genetic thompson sampling. In *2022 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8.

[Marie, 2019] Marie, F. (2019). Coliseum-3d. une plate-forme innovante pour la segmentation d'images médicales par raisonnement à partir de cas (ràpc) et méthodes d'apprentissage de type deep learning.

[Muangprathub et al., 2020] Muangprathub, J., Boonjing, V., and Chamnongthai, K. (2020). Learning recommendation with formal concept analysis for intelligent tutoring system. *Heliyon*, 6(10) :e05227.

[Nkambou et al., 2010] Nkambou, R., Bourdeau, J., and Mizoguchi, R. (2010). *Advances in Intelligent Tutoring Systems*. Springer Berlin, Heidelberg, 1 edition.

[Richter and Weber, 2013] Richter, M. and Weber, R. (2013). *Case-Based Reasoning (A Textbook)*. Springer-Verlag GmbH.

[Richter, 2009] Richter, M. M. (2009). The search for knowledge, contexts, and case-based reasoning. *Engineering Applications of Artificial Intelligence*, 22(1) :3–9.

[Robertson and Watson, 2014] Robertson, G. and Watson, I. D. (2014). A review of real-time strategy game ai. *AI Mag.*, 35 :75–104.

[Roldan Reyes et al., 2015] Roldan Reyes, E., Negny, S., Cortes Robles, G., and Le Lann, J. (2015). Improvement of online adaptation knowledge acquisition and reuse in case-based reasoning : Application to process engineering design. *Engineering Applications of Artificial Intelligence*, 41 :1–16.

[Seznec et al., 2020] Seznec, J., Menard, P., Lazaric, A., and Valko, M. (2020). A single algorithm for both restless and rested rotting bandits. In Chiappa, S. and Calandra, R., editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 3784–3794. PMLR.

[Sinaga and Yang, 2020] Sinaga, K. P. and Yang, M.-S. (2020). Unsupervised k-means clustering algorithm. *IEEE Access*, 8 :80716–80727.

[Su et al., 2022] Su, Y., Cheng, Z., Wu, J., Dong, Y., Huang, Z., Wu, L., Chen, E., Wang, S., and Xie, F. (2022). Graph-based cognitive diagnosis for intelligent tutoring systems. *Knowledge-Based Systems*, 253 :109547.

[Wang et al., 2021] Wang, F., Liao, F., Li, Y., and Wang, H. (2021). A new prediction strategy for dynamic multi-objective optimization using gaussian mixture model. *Information Sciences*, 580 :331–351.

[Xu et al., 2021] Xu, S., Cai, W., Xia, H., Liu, B., and Xu, J. (2021). Dynamic metric accelerated method for fuzzy clustering. *IEEE Access*, 9 :166838–166854.

[Zhang. and Aslan, 2021] Zhang., K. and Aslan, A. B. (2021). Ai technologies for education : Recent research and future directions. *Computers and Education : Artificial Intelligence*, 2 :100025.

[Zhao et al., 2023] Zhao, L.-T., Wang, D.-S., Liang, F.-Y., and Chen, J. (2023). A recommendation system for effective learning strategies : An integrated approach using context-dependent dea. *Expert Systems with Applications*, 211 :118535.

[Zhou and Wang, 2021] Zhou, L. and Wang, C. (2021). Research on recommendation of personalized exercises in english learning based on data mining. *Scientific Programming*, 2021 :5042286.

Document réalisé avec LaTeX et :
le style LaTeX pour Thèse de Doctorat créé par S. Galland — https://github.com/gallandarakhneorg/tex-templates
la collection de paquets `tex-upmethodology` — http://www.arakhne.org/tex-upmethodology/

école doctorale **sciences pour l'ingénieur et microtechniques**

Titre : Adaptation en temps réel d'une séance d'entraînement par intelligence artificielle

Mots-clés : Mot-clé 1, Mot-clé 2

Résumé :

L'objectif de ce travail de thèse est la prise en compte en temps réel du travail d'un apprenant pour une meilleure personnalisation d'AI-VT (Artificial Intelligence Virtual Trainer).

Title: Adaptation en temps réel d'une séance d'entraînement par intelligence artificielle

Keywords: Keyword 1, Keyword 2

Abstract:

This is the abstract in English.

Université Bourgogne-Franche-Comté
32, avenue de l'Observatoire
25000 Besançon, France