
Recommandation basée sur l'aide multicritère à la décision pour personnaliser l'échange d'information

Sarra Bouzayane^{1,3}, Inès Saad^{1,2}, Gilles Kassel¹, Faiez Gargouri³

1. Université de Picardie Jules Verne, 33 rue Saint-Leu, 80039 Amiens, France
{sarra.bouzayane,gilles.kassel,ines.saad}@u-picardie.fr

2. École Supérieure de Commerce, 18 place Saint-Michel, 80038 Amiens, France
ines.saad@esc-amiens.com

3. Institut Supérieur d'Informatique et de Multimédia, B.P. 242, Sfax 3021, Tunisie
faiez.gargouri@isims.usf.tn

RÉSUMÉ. Ce travail vise à aider les utilisateurs à trouver l'information pertinente qu'ils cherchent quand ils se trouvent face à une quantité grandissante de données numérisées produites par un nombre important d'acteurs distants et de profils hétérogènes. Nous proposons de recommander à chacun d'entre eux, une liste personnalisée d'« Apprenants leaders » susceptibles de l'accompagner durant son processus d'apprentissage médiatisé. L'approche repose sur une phase de prédiction périodique et incrémentale des « Apprenants leaders » basée sur l'aide à la décision multicritère et une phase de recommandation basée sur le filtrage démographique. Elle est validée dans un contexte des MOOC (Massive Open Online Courses).

ABSTRACT. The purpose of this work is to help the users find the relevant information they need when they are faced with a growing mass of digitized data produced by a massive number of remote actors with heterogeneous profiles. Thus, we propose to recommend to each of them a personalized list of "Leader learners" who can support them during their mediated learning process. The approach relies on a periodic and incremental prediction phase of "Leader learners" based on multicriteria decision making and a recommendation phase based on the demographic filtering. It is validated in a context of MOOCs (Massive Open Online Courses).

MOTS-CLÉS : systèmes de recommandation, échange d'informations, processus d'accompagnement, transfert des savoirs, apprenant leader, MOOC.

KEYWORDS: recommender system, information exchange, support process, knowledge transfer, leader learner, MOOC.

DOI:10.3166/ISI.22.6.71-91 © 2017 Lavoisier

1. Introduction

À l'ère du numérique, les processus d'apprentissage au sein des organisations ou bien des domaines éducatifs sont devenus de plus en plus médiatisés et centrés apprenant. L'échange de l'information se fait ainsi via un système d'information numérisé qui encourage la participation active des utilisateurs. Ces derniers peuvent enrichir le système en déposant des données numérisées, puis effectuer une recherche autonome de l'information dont ils ont besoin pour atteindre leurs objectifs. Ce nouvel environnement remet en cause les méthodes classiques de traitement de l'information qui sont devenues insuffisantes voire inefficaces et conduit à l'innovation d'autres technologies pour faire face à la massification et à l'hétérogénéité des données.

Dans ce papier, nous traitons le cas d'échange et de traitement d'information dans le contexte des MOOC (*Massive Open Online Courses*). Ce sont des systèmes d'information numérisés dédiés pour l'apprentissage en ligne et ouvert. Les MOOC sont accessibles par un nombre massif d'apprenants de profils hétérogènes. Une masse énorme de données est déposée sur le système soit par l'équipe pédagogique, soit volontairement par les apprenants sous plusieurs formes (pdf, image, vidéo, lien hypertexte). Ces données seront, ensuite, consultées par les apprenants qui fouillent dedans afin de les interpréter en informations. Ces informations seront absorbées afin d'en inférer des connaissances. Les apprenants fournissent périodiquement un retour sur les connaissances acquises en répondant sur les modalités d'évaluation proposées par l'équipe pédagogique telles que les tests automatisés et les évaluations par les pairs (cf. figure 1).

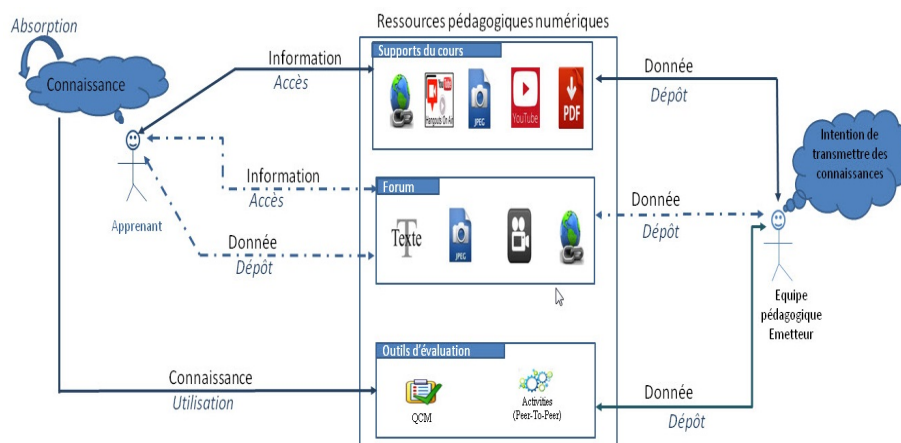


Figure 1. Système d'Information numérique dédié aux MOOC

Le MOOC est toutefois animé par une équipe pédagogique de taille réduite qui se trouve généralement incapable d'accompagner l'intégralité des participants. Ainsi, les apprenants ont recours aux données déposées sur le forum d'échange dont l'exactitude et la pertinence ne sont pas garanties. D'où notre objectif est d'identifier, parmi ce

nombre massif d'apprenants, ceux qui sont susceptibles de partager une information correcte et immédiate avec tout apprenant en besoin. Nous appelons ces apprenants « Apprenants leaders ».

Pour ce faire, nous proposons un système pour la recommandation d'une liste personnalisée des « Apprenants leaders » pour chaque apprenant en besoin en prenant en compte leurs données démographiques. En effet, la communication dans un contexte d'apprentissage médiatisé, tel que les MOOC, permet des échanges beaucoup plus naturels sans risquer qu'un temps d'attente trop long entre la question et la réponse puisse conduire à un désinvestissement de l'apprenant par rapport à la tâche qui lui est proposée (Eastmond, 1994). Elle favorise, également, l'échange mutuel de l'information entre les apprenants ce qui permet d'atteindre une compréhension réciproque pour chaque information échangée (Devirieux, 2007). De ce fait, la recommandation des individus comme étant une source d'information peut garantir l'appropriation des savoirs auprès des apprenants cibles (« Apprenants en risque » et « Apprenants en difficulté »). L'approche de recommandation est composée de deux phases : une phase de prédiction périodique des « Apprenants leaders » et une phase de recommandation. La première phase repose sur l'aide à la décision multicritère et a pour objectif de prédire d'une manière hebdomadaire les trois classes de décision : Cl_1 des « Apprenants en risque », Cl_2 des « Apprenants en difficulté » et Cl_3 des « Apprenants leaders ». Elle est basée sur l'approche DRSA (*Dominance based Rough Set Approach*) (Greco *et al.*, 2001) et sur notre algorithme DRSA-Incremental (Bouzayane, Saad, 2017b) pour la mise à jour incrémentale des approximations de ces classes de décision en considérant les préférences de l'équipe pédagogique du MOOC. La deuxième phase exploite le résultat de la première phase afin de recommander pour chaque « Apprenant en risque » ou « Apprenant en difficulté » une liste d'« Apprenants leaders » appropriée à son profil. La recommandation est basée sur le filtrage démographique et doit améliorer le processus de transfert (transmission et appropriation) de connaissances et d'informations entre l'« Apprenant leader » et l'« Apprenant en risque » ou l'« Apprenant en difficulté ».

Le papier est structuré comme suit : la section 2 présente un état de l'art sur les systèmes de recommandation proposés dans un contexte des MOOC. La section 3 présente l'approche de recommandation des « Apprenants leaders » en détaillant la phase de prédiction et celle de recommandation. La section 4 analyse les résultats des expérimentations effectuées. La section 5 résume le papier et introduit nos travaux futurs.

2. Travaux antérieurs

Etant donnée la primeur du domaine des MOOC, les systèmes de recommandation qui y sont impliqués ne sont pas nombreux. Cependant, nous pouvons les classer en deux catégories en fonction de leur objectif : les systèmes de recommandation dont l'objectif est d'aider les apprenants à choisir un MOOC approprié à leurs besoins afin de remédier au nombre massif des MOOC (Bousbahi, Chorfi, 2015 ; Symeonidis,

Malakoudis, 2016; Gutiérrez-Rojas *et al.*, 2014), et les systèmes de recommandation dont l'objectif est de supporter les apprenants durant leur processus d'apprentissage pour remédier à la massification des données échangées au sein du même MOOC. Dans ce papier, nous focalisons sur la deuxième catégorie.

Yang *et al.* (2014) ont abordé la problématique de surcharge des discussions dans les forums du MOOC qui empêche l'apprenant de trouver l'information qui l'intéresse. Ils ont proposé un modèle pour la recommandation d'une liste de discussions à chaque apprenant basé sur la correspondance entre ses préférences et le contenu de chaque discussion. Les deux filtrages collaboratif et basé-contenu sont mis en place. La recommandation est le résultat de trois modélisations : la modélisation de contenu des discussions des forums qui repose sur l'analyse des mots, la modélisation de préférences de l'apprenant qui sont extraites de l'historique de sa participation dans les discussions précédentes et la modélisation de la connexion sociale de l'apprenant basée sur ses interactions avec les pairs. Ces trois modélisations sont données en input à une matrice de factorisation adaptative (Chen *et al.*, 2012) selon une fenêtre de temps fixée afin de prédire le comportement des apprenants dans la fenêtre de temps suivante en fonction de leur comportement durant la fenêtre du temps courante. Les expérimentations ont montré qu'un tel modèle a permis d'améliorer la performance des apprenants surtout dans le cas où la fenêtre de temps choisie est réduite.

Li et Mitros (2015) ont proposé un système XBlock¹ pour recommander aux apprenants une liste de ressources de réhabilitation pertinentes pour un problème donné afin de leur fournir des interventions avec plus de profondeur et moins d'échafaudages que les forums. Le système est basé sur la technique de *crowdsourcing* où une combinaison, entre ce que les apprenants experts ont publié pour résoudre un problème et de ce que les apprenants novices ont besoin sur un thème commun, est effectuée. La *crowdsourcing* est une technique de production participative qui sollicite le grand public ou bien les consommateurs d'un système pour résoudre un problème. Dans ce contexte, elle est considérée comme une solution plus économique et plus pratique pour créer des outils de réhabilitation. Pour chaque ensemble de problème est attribué un thème et pour chaque ressource est accordé un titre, un lien, un résumé, une capture d'écran et une liste des votes. Les ressources recommandées peuvent être votées par les apprenants. L'équipe pédagogique du MOOC a le droit de modifier, supprimer ou enrichir la ressource.

Onah et Sinclair (2015) ont proposé un algorithme basé sur la technique de filtrage collaboratif pour recommander à un apprenant cible des ressources appropriées à son profil afin de supporter son parcours d'apprentissage durant sa participation au MOOC. Tout apprenant dans le système est demandé de noter chaque ressource utilisée sur une échelle de 1 à 5. Sur la base de cette notation, une fonction de prédiction est calculée afin de prédire le degré d'appréciation de l'apprenant cible à chaque ressource. Les ressources seront ainsi recommandées en fonction du degré d'appréciation prédit.

1. <https://github.com/pmitros/RecommenderXBlock>

Labarthe *et al.* (2016a) ont proposé un module de recommandation intégré pour recommander à chaque apprenant une liste d'apprenants pertinents, disponibles et prêts à partager leurs connaissances. L'objectif du travail est de favoriser l'interaction entre les apprenants. Ces derniers sont recommandés en fonction de leurs profils et activités. En cliquant sur un apprenant recommandé, l'apprenant cible aura le choix de lui envoyer un message privé, le signaler comme favorisé, le signaler comme ignoré, ou bien ouvrir une fenêtre de chat avec lui. Contrairement à un apprenant favorisé, un apprenant ignoré ne sera plus recommandé à l'apprenant cible concerné. Validé sur un MOOC de gestion des projets (Labarthe *et al.*, 2016b), ce système de recommandation a prouvé un impact considérable sur la participation des apprenants, leur complétude du MOOC, leur succès aux évaluations et leur interaction avec les autres apprenants.

Ces travaux présentent deux limites majeures. D'une part, la recommandation est destinée, particulièrement, aux participants du forum et basée sur les informations qu'ils partagent alors que ces participants ne représentent qu'une minorité limitée entre 5 % et 10 % des apprenants d'un MOOC (Kop *et al.*, 2011 ; Manning, Sanders, 2013 ; Breslow *et al.*, 2013 ; Kloft *et al.*, 2014 ; Rosé, Siemens, 2014). D'autre part, la recommandation proposée se limite à mettre à la disposition des apprenants les ressources pédagogiques (exp. une vidéo, image, pdf, etc.) qui sont appropriées à leurs profils sans s'assurer que la ressource recommandée soit, par la suite, correctement interprétée par l'apprenant récepteur.

Ainsi, notre solution doit concerner l'intégralité des apprenants d'un MOOC et réduire les inhibiteurs du processus du transfert de savoirs entre eux afin de garantir l'appropriation des informations échangées. Pour ce faire, nous recommandons une liste des individus (*i.e.* une ressource humaine) pour favoriser l'échange mutuel de l'information entre les apprenants afin d'atteindre une compréhension réciproque pour chaque information échangée (Devirieux, 2007). Comparé au travail du Labarthe *et al.* (2016a) où la sélection des apprenants pertinents était intuitive, notre système de recommandation repose sur une méthode de classification incrémentale et périodique MAI2P (*Multicriteria Approach for the Incremental Periodic Prediction*) basée sur l'aide multicritère à la décision permettant l'identification des « Apprenants leaders » en fonction des préférences de l'équipe pédagogique.

3. KTI-MOOC : un système de recommandation pour l'aide au transfert des savoirs dans les MOOC

Dans cette section nous présentons le système de recommandation KTI-MOOC (*Recommender system for the Knowledge Transfer Improvement within a MOOC*). Ce système est appliqué dans un contexte des MOOC et a l'objectif de recommander une liste ordonnée des « Apprenants leaders » qui sont susceptibles d'accompagner les « Apprenants en risque » et les « Apprenants en difficulté » tout au long de leur processus d'apprentissage. Nous présentons, dans un premier temps, le contexte de la recommandation, ensuite nous expliquons brièvement notre méthode MAI2P, et enfin nous détaillons le processus de recommandation.

3.1. Définition du contexte de la recommandation

Tout système de recommandation est caractérisé par trois éléments principaux : l'objectif de la recommandation, la cible (utilisateur courant, utilisateur actif) à qui recommander et l'objet (item, produit, service, article) à recommander.

3.1.1. Objectif de la recommandation

L'objectif de notre système de recommandation KTI-MOOC est l'aide au transfert des savoirs factuels et des savoir-faire auprès des apprenants lors de la diffusion d'un MOOC. En effet, le nombre massif d'apprenants d'une part et la taille limitée de l'équipe pédagogique d'autre part, rendent inefficace le processus de transfert et, aussi, d'accompagnement ce qui conduit les participants à abandonner la formation. En outre, l'apprentissage dans un tel contexte repose considérablement sur les informations transmises par les apprenants dont l'exactitude et la pertinence ne sont pas garanties. L'objectif de KTI-MOOC est, ainsi, de proposer une solution pour minimiser les obstacles de transfert et de permettre aux apprenants de profiter d'avantage de la formation qu'ils suivent.

3.1.2. Cible de la recommandation

L'utilisateur cible d'un système de recommandation est l'utilisateur bénéficiaire de la recommandation. Dans notre système KTI-MOOC, l'utilisateur cible est un apprenant appartenant à la classe de décision Cl_1 des « Apprenants en risque » ou bien à la classe de décision Cl_2 des « Apprenants en difficulté ». Ce sont les apprenants que nous souhaitons aider en leur permettant d'une source pertinente de l'information dont ils ont besoin afin de ne pas abandonner le MOOC. Cette liste d'apprenants varie d'une semaine à une autre durant la diffusion du MOOC. Elle est le résultat de l'application hebdomadaire de notre méthode de prédiction multicritère et incrémentale, MAI2P, dans le contexte des MOOC (cf. section 3.2).

3.1.3. Objet de la recommandation

L'objet de la recommandation est l'item à recommander à l'utilisateur cible afin de réaliser un objectif défini. Généralement, l'objet de la recommandation est un ensemble différent de l'ensemble d'utilisateurs tels que des films, des produits, des livres, des localisations, etc. Cependant, pour le cas du système de recommandation KTI-MOOC les utilisateurs cibles de la recommandation et les objets de la recommandation font partie du même ensemble. Nous recommandons des apprenants à des apprenants. L'objet de la recommandation est un ou plusieurs apprenants appartenant à la classe de décision Cl_3 des « Apprenants leaders ». De même, cet ensemble varie d'une semaine à une autre en fonction du résultat de notre méthode de prédiction MAI2P.

3.2. MAI2P : Méthode basée sur l'aide multicritère à la décision pour la prédiction périodique et incrémentale des profils apprenants

Notre méthode MAI2P a l'objectif de prédire les « Apprenants en risque », les « Apprenants en difficulté » et les « Apprenants leaders » pendant la semaine suivante du MOOC en fonction de leurs données statiques et dynamiques de la semaine en cours. Cette méthode est basée sur une approche d'aide à la décision multicritère DRSA (*Dominance-based Rough Set Approach*) (Greco *et al.*, 2001) et sur un algorithme DRSA-Incremental (Bouzayane, Saad, 2017b) que nous avons proposé pour la mise à jour incrémentale des approximations de DRSA suite à l'addition d'un ensemble d'objets dans le système d'information. La méthode MAI2P est composée de trois phases telles que la première et la deuxième phase sont exécutées à la fin de chaque semaine S_i du MOOC alors que la troisième phase est exécutée au début de chaque semaine S_{i+1} du même MOOC avec $i \in \{1..t-1\}$, où t est la durée du MOOC en semaines.

La première phase repose, d'abord, sur la construction d'une famille F_i de p critères permettant de caractériser un profil apprenant (par exemple le niveau d'études, la motivation d'inscription, la fréquence d'accès au site du MOOC, etc.). Chaque critère doit permettre de mesurer les préférences de l'équipe pédagogique du MOOC selon un point de vue personnel (Roy, 1997). Prenons l'exemple du critère « Niveau d'études », les préférences que nous pouvons y appliquer sont : 1 : élève ; 2 : étudiant ; 3 : doctorant ; 4 : docteur. L'étape de construction d'une famille de critères est détaillée dans (Bouzayane, Saad, 2016). Ensuite, à chaque semaine S_i du MOOC l'équipe pédagogique doit construire un échantillon d'apprentissage K_i contenant un ensemble de m apprenants, représentatif pour chacune des trois classes de décision prédéfinies :

- Cl_1 . La classe de décision des « Apprenants en risque » contenant ceux qui risquent d'abandonner le cours au bout de la semaine suivante du MOOC.
- Cl_2 . La classe de décision des « Apprenants en difficulté » définie par ceux qui ont des difficultés mais qui sont toujours actifs et n'ont pas l'intention de quitter le MOOC au moins pour la semaine suivante.
- Cl_3 . La classe de décision des « Apprenants leaders » représentée par ceux qui sont capables de diriger une équipe d'apprenants en leur fournissant une réponse précise et immédiate.

A chaque apprenant $App_{i,j}$ tel que $j \in \{1..m\}$, contenu dans l'échantillon d'apprentissage K_i sélectionné dans la semaine S_i , sera attribué un vecteur d'évaluations sur l'ensemble de p critères en fonction de l'ordre de préférence préfixé. Ce vecteur d'évaluation va permettre à l'équipe pédagogique de classer chaque apprenant de l'échantillon d'apprentissage dans l'une des trois classes de décision Cl_1 , Cl_2 ou Cl_3 .

La deuxième phase est basée sur un algorithme de mise à jour incrémentale des approximations inférieures et supérieures de l'approche DRSA suite à l'addition d'un ensemble d'apprenants à l'échantillon d'apprentissage. En effet, vu l'entrée/sortie libre des apprenants du MOOC, l'échantillon d'apprentissage ne peut pas être stable d'une semaine à une autre. Il faut donc sélectionner pour chaque semaine S_i un échantillon

d'apprentissage K'_i qui s'ajoute à l'ensemble K_{i-1} tel que $K_i = K_{i-1} + K'_i$. Ainsi, pour éviter de recalculer les approximations de la méthode de DRSA de zéro, nous avons proposé un algorithme pour leur mise à jour incrémentale. Ces approximations seront données en entrée à un algorithme d'induction pour l'inférence d'un ensemble des règles de décision sous la forme de « Si... Alors... ».

La dernière phase consiste à l'exploitation des règles de décision inférées à la fin de la semaine S_i afin d'assigner, au début de la semaine S_{i+1} , les apprenants potentiels dans l'une des trois classes de décision Cl_1 des « Apprenants en risque », Cl_2 des « Apprenants en difficulté » ou Cl_3 des « Apprenants leaders ». Nous entendons par apprenants potentiels ceux qui sont susceptibles d'être classés dans l'une des trois classes de décision. Notre méthode MAI2P est détaillée dans (Bouzayane, Saad, 2017a ; 2017c). Nous rappelons que la liste hebdomadaire des « Apprenants en risque » et des « Apprenants en difficulté », prédite par la méthode MAI2P, représente l'ensemble des utilisateurs cibles du système KTI-MOOC et celle des « Apprenants leaders » représente l'objet de la recommandation.

3.3. Processus de recommandation des « Apprenants leaders »

Le processus de recommandation adopté par le système KTI-MOOC repose sur la technique de filtrage démographique. Ce filtrage catégorise les utilisateurs en fonction de leurs données démographiques (exp. sexe, âge, pays, niveau d'études, etc.). Les recommandations sont produites selon les évaluations des utilisateurs du même groupe (appelé aussi niche ou voisinage) sur l'ensemble des items. Le filtrage démographique repose sur l'hypothèse que deux utilisateurs qui ont évolué dans un même environnement sont plus susceptibles d'avoir le même goût et les mêmes préférences. Le processus de recommandation est composé de trois étapes : 1) le profilage, 2) le calcul du voisinage en utilisant une fonction de similarité, et 3) la prédiction des appréciations des utilisateurs (cf. figure 2).

3.3.1. Modélisation du profil apprenant

La modélisation du profil utilisateur est un domaine de recherche qui concerne l'amélioration de l'interaction homme-machine en prédisant les intérêts des utilisateurs (Shani *et al.*, 2007). Le profil repose sur deux notions clés qui sont le modèle de représentation et les informations à considérer. Dans la littérature, plusieurs modèles de représentation ont été proposés, dont le modèle vectoriel (Dumais *et al.*, 2016), le modèle sémantique (Sridevi, Umarani, 2013) et le modèle multidimensionnel (Amato, Straccia, 1999). Dans ce travail, nous adoptons la représentation vectorielle puisque les informations que nous traitons font partie de la même dimension et sont considérées comme sémantiquement indépendantes. Les informations à inclure dans le modèle de représentation peuvent être entrées manuellement par l'utilisateur, ou extraites automatiquement à partir du contenu des ressources qu'il a consultées (Bonnin, 2010). Elles doivent satisfaire la finalité de la recommandation qui est dans notre cas l'aide au transfert des savoirs factuels et du savoir-faire. Ces informations doivent permettre au système de recommandation de minimiser la distance entre le profil de l'apprenant

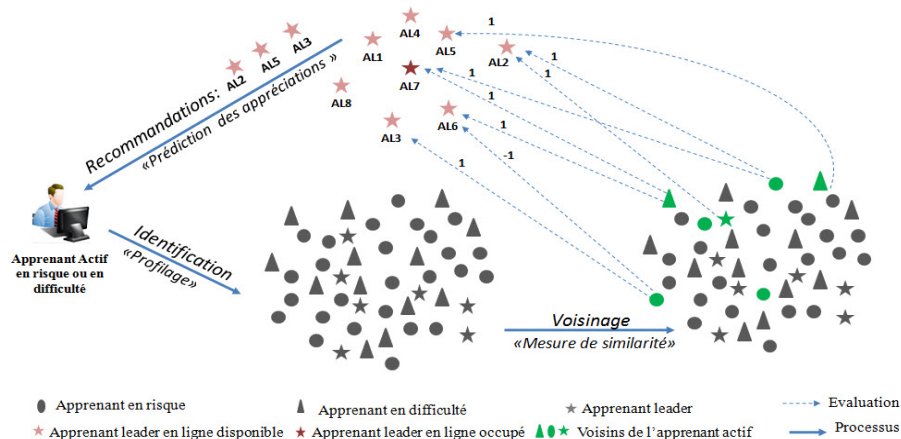


Figure 2. Processus de recommandation démographique : cas d'un « Apprenant en risque » ou un « Apprenant en difficulté »

transmetteur des savoirs et celui de l'apprenant récepteur de ces savoirs afin de mener au mieux le processus de transfert entre eux. En d'autres termes, ces informations doivent représenter les facteurs inhibant le processus de transfert tels que la langue, le pays et le domaine d'études.

- La langue : la langue est un obstacle majeur au processus de transfert de savoirs. Dans le domaine de gestion des connaissances, Welch et Welch (2008) ont montré que la barrière de la langue impacte la capacité de transmettre la connaissance de la part du transmetteur, d'une part, et la capacité de l'absorber auprès du récepteur, d'autre part. L'étude effectuée par Chiu *et al.* (2006) a aussi prouvé l'impacte de la langue sur la qualité de la connaissance transférée. Dans le contexte du MOOC, Barak (2015) a prouvé que la langue est une forte barrière pour la compréhension et la complétude du MOOC.

- Le domaine d'études : le domaine d'études partagé permet aux acteurs d'avoir un langage scientifique et technique similaire. Selon (Grundstein, 2009), les personnes qui partagent la même culture peuvent avoir des schémas d'interprétation (Arduin, 2013) proches leur permettant de donner un même sens à une connaissance codifiée, ce qui aide à augmenter le niveau de la commensurabilité (Kuhn *et al.*, 1970) entre eux et donc à améliorer le processus de transfert. Gooderham (2007) a aussi mis l'accent sur l'impact de la distance culturelle sur la qualité du processus de transfert de connaissances. Cela dit que les personnes issues du même domaine d'étude peuvent se comprendre mieux que deux personnes de cultures différentes.

- La distance géographique : par rapport à l'interaction en face-à-face, l'interaction à distance met beaucoup d'inconvénients surtout lorsqu'il s'agit d'un transfert du savoir-faire. La distance géographique a été prouvée par Gooderham (2007) comme étant un inhibiteur du processus de transfert des connaissances. Ambos et Ambos (2009) ont constaté que plus la distance entre le transmetteur et le récepteur est

courte, plus l'efficacité du transfert des connaissances est élevée. Dans le contexte des MOOC, cela permet aussi de prendre en considération le décalage entre les faisceaux horaires qui complique la correspondance entre les apprenants. Il est préférable que le transmetteur et le récepteur proviennent de la même ville pour tenter une interaction en face-à-face si possible.

Ces informations sont à considérer pour construire les profils des apprenants et, par suite, calculer la similarité entre les « Apprenants leaders » et les apprenants cibles. Il est préférable de minimiser la distance linguistique, culturelle et géographique entre l'apprenant transmetteur et l'apprenant récepteur des savoirs.

3.3.2. Identification du voisinage

Le voisinage d'un apprenant cible est représenté par les apprenants qui lui sont plus proches considérant la langue, le domaine d'études et la distance géographique. Nous sommes donc face à un problème de minimisation de distance. Plusieurs fonctions de similarité ont été définies dans la littérature dont les plus prometteuses sont le coefficient de corrélation de Pearson, la similarité de cosinus et la distance euclidienne.

La similarité de Cosinus s'applique uniquement sur des valeurs numériques et permet de calculer la similarité entre deux vecteurs en déterminant le cosinus de l'angle entre eux. Elle est fréquemment utilisée en tant que mesure de ressemblance entre deux documents et elle exige donc un moyen de convertir les mots d'un document en nombres (par exemple chercher un document vectorisé dans un corpus de documents aussi vectorisés). Le coefficient de corrélation de Pearson s'applique sur deux variables dépendantes de caractère quantitatif dont l'une permet d'expliquer l'autre (par exemple la corrélation entre l'altitude en centaines de mètres et la température au niveau de la mer). Ces deux mesures ne sont pas donc applicables dans notre contexte ce qui justifie notre choix de la distance euclidienne.

La distance euclidienne a une limite inférieure égale à zéro indiquant une correspondance parfaite sans limite supérieure proportionnelle. Soient $(x_1, x_2, \dots, x_k, \dots, x_n)$ et $(y_1, y_2, \dots, y_k, \dots, y_n)$ les représentations vectorielles relatives aux profils de deux apprenants x et y respectivement. La distance euclidienne entre les deux profils est calculée comme suit :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

tel que $x_i - y_i = 0$ si $x_i = y_i$ et $x_i - y_i = 1$ si $x_i \neq y_i$.

Dans notre cas nous considérons uniquement quatre attributs ($n = 4$) qui sont la langue, le domaine d'études, le pays et la ville d'où la représentation d'un apprenant X est comme suit :

$$X = (x_1 = \text{Langue}, x_2 = \text{Domaine d'études}, x_3 = \text{Pays}, x_4 = \text{Ville})$$

Aussi, nous considérons que les quatre attributs ont la même importance et donc aucune pondération ne sera appliquée.

3.3.3. Prédiction basée sur le filtrage démographique

Comme nous l'avons déjà expliqué, le filtrage démographique est basé sur les évaluations faites par le voisinage démographique de l'utilisateur cible. La recommandation, dans notre cas, est la liste des « Apprenants leaders » qui ont été précédemment appréciés par les apprenants similaires à l'apprenant cible.

Afin de recommander à un apprenant la liste de leaders appropriés à son profil et qui sont susceptibles de l'accompagner durant une session de navigation sur l'environnement du MOOC, nous devons prédire le taux d'appréciation $\hat{r}_{c,l}$ d'un apprenant cible c pour un apprenant leader l , en utilisant les évaluations données par son voisinage pour ce même apprenant leader. Plusieurs fonctions d'agrégation ont été proposées dans la littérature pour la mesure d'appréciation dans les deux approches basée-contenu et collaborative (Ricci *et al.*, 2011 ; Adomavicius, Tuzhilin, 2005). Dans ce travail, nous nous intéressons aux approches collaboratives et nous adaptons la fonction d'agrégation la plus courante qui consiste à utiliser la somme pondérée comme suit :

$$\hat{r}_{c,l} = \frac{\sum_{(v \in V_l(c))} w_{c,v} r_{v,l}}{\sum_{(v \in V_l(c))} w_{c,v}} \quad (2)$$

Dans l'équation (2), les variables c, l, v désignent respectivement l'apprenant cible, l'apprenant leader et l'apprenant voisin de l'apprenant cible. L'ensemble $v_l(c)$ est le voisinage de l'apprenant cible qui a évalué l'« Apprenant leader » en question. La variable $w_{c,v}$ reflète le poids du voisin, calculé par son degré de similarité avec l'apprenant cible. Le taux $r_{v,l}$ est l'évaluation donnée par le voisin v à l'« Apprenant leader » l en question.

Le dénominateur de l'équation (2) a été ajouté dans un objectif de normalisation et éviter le cas où la somme des poids dépasse la valeur 1 qui peut donner une valeur d'appréciation prédite hors intervalle. Cette équation permet également à l'utilisateur d'utiliser sa propre fonction de similarité puisque la somme sera toujours normalisée. Ainsi, nous considérons que cette formule est appropriée pour calculer le taux d'appréciation dans notre domaine d'application. Pour plus de détails sur cette mesure et les autres voir (Ricci *et al.*, 2011).

Afin de prendre en considération les cas particuliers possibles et de gérer les situations de conflits, nous avons appliqué les hypothèses simplificatrices suivantes :

- HS1 : Afin de faire face au nombre des « Apprenants en risque » et des « Apprenants en difficulté » élevé par rapport au nombre des « Apprenants leaders » en ligne, nous fixons le nombre des « Apprenants leaders » recommandés à trois. Autrement dit, un apprenant cible ne peut pas contacter plus que trois « Apprenants leaders » à la fois.

– HS2 : Afin de respecter la capacité humaine d'un « Apprenant leader » nous lui proposons un maximum d'apprenants cibles en risque ou en difficulté qui ne dépasse pas trois à la fois. Si l'un des apprenants cibles, avec qui l'« Apprenant leader » est en train d'échanger, se déconnecte, il sera remplacé par un autre apprenant à condition que le nombre total reste égal à trois.

– HS3 : Un « Apprenant leader » évalué comme « non pertinent » par un apprenant cible ne lui sera plus recommandé même s'il a été évalué pertinent par son voisinage.

– HS4 : Un « Apprenant leader » apprécié par un apprenant cible, lui sera automatiquement recommandé à l'entête de la liste dès son identification à condition que cet « Apprenant leader » soit en ligne et aussi disponible, donc en train d'échanger avec moins de trois personnes. Cette stratégie du premier arrivé premier servi, permet d'éviter le conflit dans le cas d'un « Apprenant leader » apprécié par deux apprenants cibles en besoin.

– HS5 : En cas de conflit entre un « Apprenant en risque » et un « Apprenant en difficulté », nous donnons priorité à celui qui est en difficulté partant de l'idée qu'il est le plus motivé à compléter le MOOC. Puisqu'un « Apprenant en risque » est plus susceptible d'abandonner la formation, il est plus profitable d'accompagner un apprenant en difficulté. Rappelons que la motivation est un critère qui a été considéré lors de la prédiction de la classe de décision à laquelle appartiendra un apprenant.

– HS6 : En cas d'un nouvel « Apprenant leader » qui n'est pas encore évalué ou bien le cas de manque des « Apprenants leaders » disponibles, le système complète la liste de trois « Apprenants leaders » à recommander à l'apprenant cible par des « Apprenants leaders » de son voisinage qui soient en ligne et disponibles. Dans ce cas, le système recommande à l'apprenant cible les « Apprenants leaders » qui lui sont similaires au lieu des « Apprenants leaders » appréciés par les apprenants qui lui sont similaires. De même, si un apprenant s'inscrit au milieu de la semaine, et donc n'est pas encore affecté à une classe de décision, il sera automatiquement classé dans la classe Cl_2 des « Apprenants en difficulté ».

L'algorithme de recommandation des « Apprenants leaders » doit prendre en considération ces hypothèses simplificatrices. Les « Apprenants leaders » en ligne disponibles et d'une valeur $\hat{r}_{c,l}$ la plus élevée seront recommandés et affichés sur la page personnelle de l'apprenant cible. La liste de trois peut être complétée par des « Apprenants leaders » en ligne disponibles de son voisinage en cas de manque (cf. figure 3).

Ces hypothèses simplificatrices permettent de surmonter quelques limites de filtrage démographique telles que le problème de démarrage à froid. Cependant, cette technique, appliquée dans le domaine des MOOC, demeure soumise à deux limitations majeures : le problème de *Gray Sheep* représenté par les apprenants atypiques qui peuvent exister et dont il est difficile d'en prédire l'appréciation. Le second problème est celui de *shillig* qui se produit lorsque les apprenants se servent des faux profils.

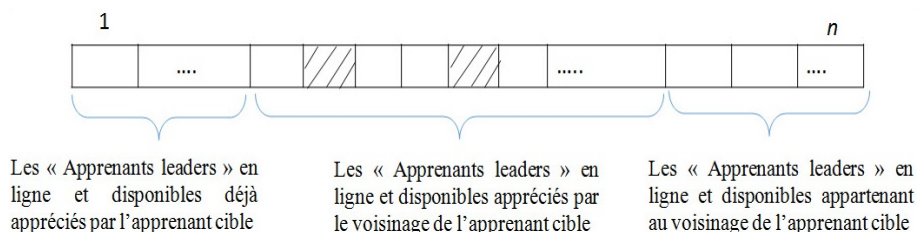


Figure 3. Répartition de n -top « Apprenants leaders » recommandés (les cases hachées représentent les « Apprenants leaders » non appréciés par l'apprenant cible)

4. Expérimentations et résultats

Dans cette section nous présentons, d'abord, quelques interfaces du système de recommandation KTI-MOOC et nous l'évaluons, ensuite, sur trois aspects : la qualité de la recommandation, la performance sur le temps d'exécution et la couverture de l'espace item (Shani, Gunawardana, 2011). Les algorithmes sont codés avec Java et exécutés sur une machine personnelle avec Windows 7, Intel (R) *CoreTM* i3-3110M CPU 2.4 GHz et 4.0 GB de mémoire.

4.1. Quelques interfaces

La figure 4 représente la page d'accueil accessible pour tout apprenant inscrit au MOOC. C'est l'interface lui permettant d'accéder à sa page personnelle une fois que les données d'identification sont valides.



Figure 4. Page d'accueil de l'environnement du MOOC

Dès que l'apprenant clique sur le bouton « Valider », l'algorithme de recommandation s'exécute afin de l'identifier (cf. figure 5). Si l'apprenant X est un « Apprenant en

risque » ou « Apprenant en difficulté », alors il sera identifié comme étant un utilisateur cible, auquel sont calculés le voisinage et la fonction de prédiction afin d'en inférer les n-top « Apprenants leaders » appropriés. Si l'apprenant X est un « Apprenant leader », il présentera l'objet de la recommandation.

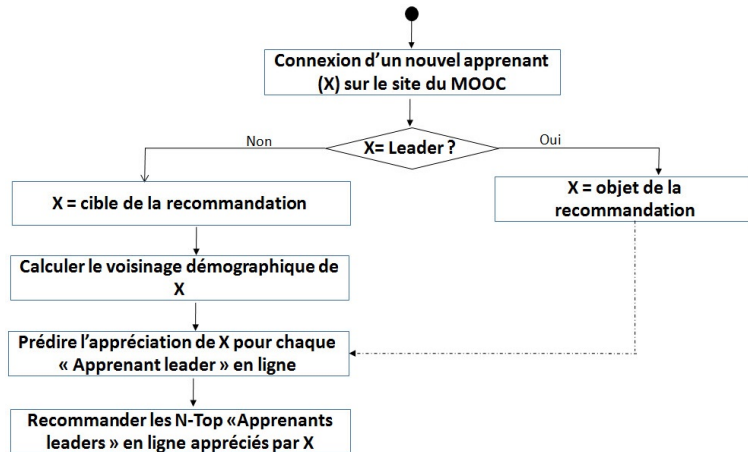


Figure 5. Processus de recommandation pour un apprenant X

La figure 6 représente la page personnelle d'un « Apprenant en risque » ou un « apprenant en difficulté ». Le système de recommandation KTI-MOOC est un module intégré dans l'environnement du MOOC qui doit apparaître sous forme d'une liste déroulante, appelée *widget*, sur la page personnelle de l'apprenant cible, en bas et à droite de la page. La liste contient l'ensemble des « Apprenants leaders » personnalisée en fonction du profil de l'apprenant cible en question.



Figure 6. Page personnelle d'un apprenant du MOOC suite à l'intégration du widget KTI-MOOC

Dès qu'il clique sur un nom de la liste recommandée, une fenêtre de *Chat* est ouverte pour l'apprenant cible (« Apprenant en risque » ou « apprenant en difficulté ») lui permettant (par ordre d'apparition sur la fenêtre de *Chat*) de :

- Ouvrir une discussion audio ou bien audio-visuelle avec l'« Apprenant leader » recommandé.
- Fournir un retour de pertinence sur l'« Apprenant leader » contacté (bouton *like* pour un retour positif et un bouton *dislike* pour un retour négatif).
- Joindre et envoyer un fichier (exp. image, pdf, vidéo, etc.).
- Ajouter un autre apprenant à la discussion.
- Fermer la fenêtre de *Chat*.

L'apprenant cible peut, sinon, se contenter d'une discussion par écrit en utilisant la fenêtre affichée.

4.2. Qualité de la recommandation

Afin d'évaluer la qualité des recommandations que le système propose, nous avons appliqué notre méthode de classification multicritère, périodique et incrémentale MAI2P sur des données réelles provenant d'un MOOC français proposé par une école de commerce en France. Ce MOOC a duré cinq semaines et accédé par 2360 apprenants. Dans ces expérimentations, uniquement 1 535 apprenants ont été considérés. Les apprenants négligés sont ceux qui n'ont pas rempli le formulaire d'inscription ce qui empêche leurs évaluations sur l'ensemble de critères.

Nous évaluerons, ainsi, la qualité de prédiction d'une semaine à une autre durant la diffusion du MOOC et pour chaque classe de décision : Cl_1 des « Apprenants en risque », Cl_2 des « Apprenants en difficulté » et Cl_3 des « Apprenants leaders ». Nous utilisons la mesure F-mesure qui est proposée pour harmoniser les deux mesures de rappel et de précision.

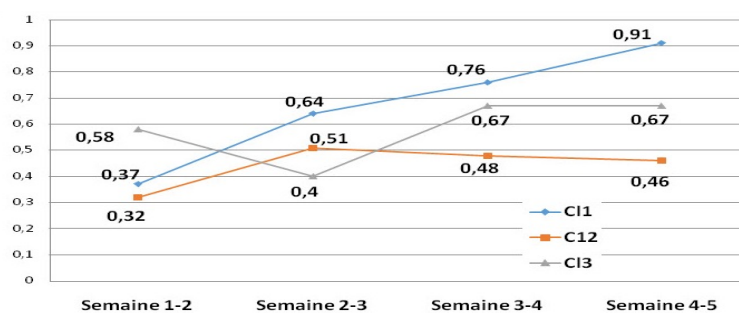


Figure 7. La variation de la F-mesure de chacune des classes de décision pendant les quatre semaines du MOOC

La figure 7 présente la variation de la F-mesure des trois classes de décision Cl_1 ,

Cl_2 et Cl_3 d'une semaine à une autre tout au long de la diffusion du MOOC. Les points clés que nous pouvons révéler et expliquer dans ces courbes sont comme suit :

- La F-mesure de la classe Cl_1 , des « Apprenants en risque », augmente au cours du temps : cela dit que la performance de la prédiction concernant la classe Cl_1 augmente au cours du temps. En effet, le MOOC est connu par les *lurkers*. Ces apprenants restent actifs au bout de la première semaine mais en ayant une intention préalable d'abandonner la formation. Ce type d'apprenants dégrade la performance du modèle de prédiction qui est basé sur le profil et le comportement de l'apprenant et non pas sur son intention. Par conséquent, plus le nombre de *lurkers* diminue plus la qualité de la prédiction est meilleure.

- La F-mesure de la classe Cl_3 , des « Apprenants leaders » augmente progressivement au cours du temps. En effet, d'une semaine à une autre, les apprenants multiplient leur participation au forum ce qui donne une information plus ample sur leurs profils. Aussi, les évaluations proposées par le MOOC sont de plus en plus complexes d'une semaine à une autre. Evidemment, une évaluation complexe (*i.e peer-to-peer* ou les activités évaluées par les pairs) permet de mieux évaluer un apprenant qu'un simple Quiz. Cela est bien justifié par la dégradation de la F-mesure de la classe Cl_3 dans la semaine 2-3. En effet, le MOOC a proposé un quiz à la fin de la semaine 2 et une activité *peer to peer* à la fin de la troisième semaine. Cependant, les apprenants qui passent un quiz avec succès peuvent se bloquer au niveau d'une activité *peer-to-peer* ou même abandonner le MOOC à cause de la complexité de cette dernière. Ceci est susceptible de se répercuter sur la qualité de la prédiction.

La qualité de prédiction est très satisfaisante qui atteint 91 % pour la classe Cl_1 et 67 % pour la classe Cl_3 . Ces mesures sont analysées avec plus de détails dans (Bouzayane, Saad, 2017c).

4.3. Performance de l'algorithme de recommandation sur le temps d'exécution

La figure 8 montre le résultat des simulations effectuées sur des données semi-réelles tout en modifiant le nombre d'apprenants et aussi le nombre des évaluations enregistrées dans la base de données.

Nous avons ainsi constaté que l'algorithme de recommandation que nous proposons est sensible aux deux facteurs suscités. L'algorithme de recommandation est plus rapide lorsque moins d'apprenants sont inscrits et moins d'évaluations sont données. Cela est lié au fait que le filtrage démographique traite les données démographiques de tous les apprenants ainsi que les évaluations qu'ils soumettent.

4.4. Évaluation sur la couverture de l'espace item

Plus que la qualité de la recommandation, il est aussi important d'évaluer la couverture de l'espace item (« Apprenants leaders ») du système proposé. Cette couverture désigne la proportion des « Apprenants leaders » que le système de recommandation peut recommander. Elle est calculée par le pourcentage des « Apprenants leaders »

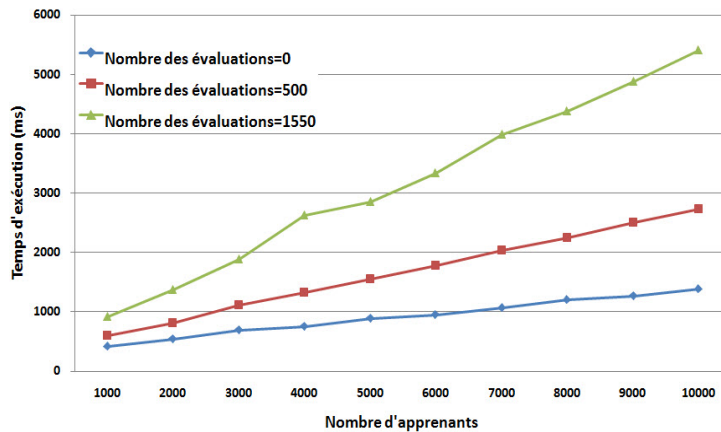


Figure 8. Temps d'exécution de l'algorithme : étude de la sensibilité par rapport au nombre d'apprenants et celui des évaluations

recommandés par le système par rapport au nombre total des « Apprenants leaders ». Dans notre cas, la couverture de l'espace item dépend du nombre des « Apprenants leaders » disponibles à recommander et aussi du nombre des « Apprenants en risque » ou bien les « Apprenants en difficulté » que nous envisageons aider simultanément.

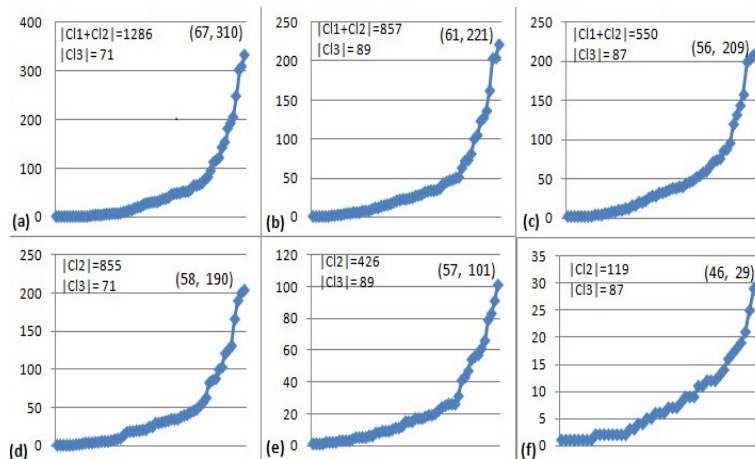


Figure 9. Couverture de l'espace item et diversité des recommandations proposées (abscisses : identifiant de l'apprenant leader recommandé ; ordonnées : nombre de recommandations d'un apprenant leader)

La figure 9 représente les résultats des simulations effectuées sur des ensembles distincts de données en modifiant la taille de l'ensemble cible de recommandation. Les « Apprenants leaders » sont ordonnés selon un ordre croissant de leur fréquence de recommandation. Dans les courbes supérieures ((a), (b) et (c)) la recommandation

concerne les « Apprenants en risque » classés dans la classe Cl_1 et les « Apprenants en difficulté » classés dans la classe de décision Cl_2 . Par contre, dans les courbes inférieures ((d), (e) et (f)) la recommandation concerne uniquement les « Apprenants en difficulté ». Nous constatons que la couverture sur l'espace item diminue en diminuant la taille de l'ensemble cible de la recommandation ($\frac{67}{71}=0,94$, $\frac{61}{89}=0,68$ et $\frac{56}{87}=0,64$ respectivement pour les courbes (a), (b) et (c); $\frac{58}{71}=0,81$, $\frac{57}{89}=0,64$ et $\frac{46}{87}=0,52$ respectivement pour les courbes (d), (e) et (f)). Cela revient à l'hétérogénéité des profils apprenants qui diminue en fonction de leur nombre et aussi à la borne que nous avons imposée sur la taille de la recommandation (3 apprenants leaders uniquement peuvent être recommandés pour chaque apprenant en besoin) qui ont empêché de trouver des paires démographiques pour quelques « Apprenants leaders » et par conséquent nous obtenons une couverture plus réduite. Dans ce cas, un « Apprenant leader » sera recommandé plusieurs fois ce qui influe la diversité de la recommandation. Cependant, dans tous les cas, plus que la moitié des « Apprenants leaders » a été recommandée.

5. Conclusion

Dans ce papier nous avons proposé notre système de recommandation KTI-MOOC conçu pour l'aide au transfert de savoirs dans un contexte des MOOC. Il a l'objectif d'aider les apprenants à trouver un apprenant source d'une information pertinente. L'approche de recommandation est basée sur deux phases : la prédiction périodique et incrémentale des « Apprenants en risque », des « Apprenants en difficulté » et des « Apprenants leaders » qui représentent, à la fois, la cible et l'objet de la recommandation. Cette phase est basée sur notre méthode MAI2P qui intègre notre algorithme DRSA-Incremental de mise à jour incrémentale des approximations de DRSA. La seconde phase consiste à appliquer un algorithme de recommandation afin de proposer pour chaque « Apprenant en risque » ou « Apprenant en difficulté », la liste des « Apprenants leaders » appropriée à son profil. Cette étape de correspondance, ou *matching*, a pris en compte les facteurs inhibiteurs d'un processus de transfert de savoirs. La recommandation est basée sur le filtrage démographique et la distance euclidienne. Des hypothèses simplificatrices ont été aussi définies pour surmonter quelques limitations.

Le système de recommandation KTI-MOOC a été également développé et validé sur des données semi-réelles sur les deux aspects : le temps d'exécution et la couverture minimale qui était d'une valeur très satisfaisante atteignant 0,94. La qualité des items recommandés (« Apprenants leaders ») a été également validée et a prouvé des résultats satisfaisants considérant la F-mesure.

Le système de recommandation KTI-MOOC est un module intégré dans l'environnement du MOOC qui doit apparaître sous forme d'une liste déroulante, appelée *widget*, sur la page personnelle de l'apprenant cible, en bas et à droite de la page. La liste contient l'ensemble des « Apprenants leaders » personnalisée en fonction du profil de l'apprenant cible en question. La fenêtre de chat permet à l'apprenant cible de déclencher une communication audio-visuelle avec l'« Apprenant leader » recommandé, évaluer son intervention et son aide ou bien lui écrire un message.

Contrairement aux travaux antérieurs qui recommandaient des ressources pédagogiques, nous avons choisi de recommander des ressources humaines. Cela permet de mener un processus de transfert où le transmetteur et le récepteur des savoirs interagissent et communiquent ensemble afin de garantir la phase d'appropriation que nous jugeons primordiale pour le processus de transfert. La communication supportée par la technologie pour l'échange de solutions peut parfaitement constituer le point de départ à un réel transfert de connaissances et donner lieu à une construction mutuelle de savoirs (Depover *et al.*, 2000).

À ce niveau, nous n'avons pas intégré notre système KTI-MOOC sur les plateformes dédiées aux MOOC afin de pouvoir valider son utilité et son effet réels sur les processus d'apprentissage des apprenants. Une telle expérimentation nous permettra, cependant, de vérifier la fréquence d'utilisation du système par les apprenants et son impact sur la qualité du processus d'apprentissage qu'ils mènent, d'une part, et la minimisation du taux d'abandon, d'autre part. Ce point sera notre objectif dans des travaux futurs.

Bibliographie

- Adomavicius G., Tuzhilin A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, vol. 17, n° 6, p. 734–749.
- Amato G., Straccia U. (1999). User profile modeling and applications to digital libraries. In *International conference on theory and practice of digital libraries*, p. 184–197.
- Ambos T. C., Ambos B. (2009). The impact of distance on knowledge transfer effectiveness in multinational corporations. *Journal of International Management*, vol. 15, n° 1, p. 1–14.
- Arduin P.-E. (2013). *Vers une métrique de la commensurabilité des schémas d'interprétation*. Thèse de doctorat non publiée, Université Paris-Dauphine.
- Barak M. (2015). The same mooc delivered in two languages: Examining knowledge construction and motivation to learn. *Proceedings of the EMOOCS*, p. 217–223.
- Bonnin G. (2010). *Vers des systèmes de recommandation robustes pour la navigation web: inspiration de la modélisation statistique du langage*. Thèse de doctorat non publiée, Université Nancy II.
- Bousbahi F., Chorfi H. (2015). Mooc-rec: a case based recommender system for moocs. *Procedia-Social and Behavioral Sciences*, vol. 195, p. 1813–1822.
- Bouzayane S., Saad I. (2016). A preference ordered classification to leader learners identification in a mooc. *Journal of Decision Systems*, p. 1–14.
- Bouzayane S., Saad I. (2017a). Dominance-based rough set approach to learn in the presence of class imbalance. *IJCAI 2017 Workshop on Learning in the Presence of Class Imbalance and Concept Drift (LPCICD'17)*, p. 24–31.
- Bouzayane S., Saad I. (2017b). Incremental updating algorithm of the approximations in drsa to deal with the dynamic information systems of moocs. In *International conference on knowledge management, information and knowledge systems (kmiks 2017)*, p. 55–66.

- Bouzayane S., Saad I. (2017c). Prediction method based drsa to improve the individual knowledge appropriation in a collaborative learning environment: case of moocs. In *Proceedings of the 50th hawaii international conference on system sciences*, p. 24–33.
- Breslow L., Pritchard D. E., DeBoer J., Stump G. S., Ho A. D., Seaton D. T. (2013). Studying learning in the worldwide classroom: Research into edx's first mooc. *Research & Practice in Assessment*, vol. 8.
- Chen T., Zhang W., Lu Q., Chen K., Zheng Z., Yu Y. (2012). Svdfeature: a toolkit for feature-based collaborative filtering. *Journal of Machine Learning Research*, vol. 13, n° Dec, p. 3619–3622.
- Chiu C.-M., Hsu M.-H., Wang E. T. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision support systems*, vol. 42, n° 3, p. 1872–1888.
- Depover C. *et al.* (2000). Un modèle d'apprentissage à distance basé sur le partage des connaissances.
- Devirieux C. J. (2007). *Pour une communication efficace: Quoi dire et comment le dire*. PUQ.
- Dumais S., Cutrell E., Cadiz J. J., Jancke G., Sarin R., Robbins D. C. (2016). Stuff i've seen: a system for personal information retrieval and re-use. In *Acm sigir forum*, vol. 49, p. 28–35.
- Eastmond D. V. (1994). Adult distance study through computer conferencing. *Distance Education*, vol. 15, n° 1, p. 128–152.
- Gooderham P. N. (2007). Enhancing knowledge transfer in multinational corporations: a dynamic capabilities driven model. *Knowledge Management Research & Practice*, vol. 5, n° 1, p. 34–43.
- Greco S., Matarazzo B., Slowinski R. (2001). Rough sets theory for multicriteria decision analysis. *EJOR*, vol. 129, n° 1, p. 1–45.
- Grundstein M. (2009). Gameth®: a constructivist and learning approach to identify and locate crucial knowledge. *International Journal of Knowledge and Learning*, vol. 5, n° 3-4, p. 289–305.
- Gutiérrez-Rojas I., Leony D., Alario-Hoyos C., Pérez-Sanagustín M., Delgado-Kloos C. (2014). Towards an outcome-based discovery and filtering of moocs using moocrank. *Proceedings of the Second MOOC European Stakeholders Summit*, p. 50–57.
- Kloft M., Stiehler F., Zheng Z., Pinkwart N. (2014). Predicting mooc dropout over weeks using machine learning methods. In *Proceedings of the emnlp 2014 workshop on analysis of large scale social interaction in moocs*, p. 60–65.
- Kop R., Fournier H., Mak J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? participant support on massive open online courses. *The International Review Of Research In Open And Distributed Learning*, vol. 12, n° 7, p. 74–93.
- Kuhn T. S., Lakatos I., Musgrave A. (1970). Criticism and the growth of knowledge. *Criticism and the growth of knowledge*.
- Labarthe H., Bachelet R., Bouchet F., Yacef K. (2016a). Increasing mooc completion rates through social interactions: a recommendation system. *EMOOCs, Research Track*, p. 471–480.

- Labarthe H., Bouchet F., Bachelet R., Yacef K. (2016b). Does a peer recommender foster students' engagement in moocs? In *9th international conference on educational data mining*, p. 418–423.
- Li S.-W. D., Mitros P. (2015). Learnersourced recommendations for remediation. In *Advanced learning technologies (icalt), 2015 IEEE 15th international conference on*, p. 411–412.
- Manning J., Sanders M. (2013). How widely used are mooc forums? a first look. *Signals: Thoughts on Online Learning*.
- Onah D. F., Sinclair J. (2015). Collaborative filtering recommendation system: a framework in massive open online courses. *INTED2015 Proceedings*, p. 1249–1257.
- Ricci F., Rokach L., Shapira B. (2011). *Introduction to recommender systems handbook*. Springer.
- Rosé C. P., Siemens G. (2014). Shared task on prediction of dropout over time in massively open online courses. In *Proc. of emnlp*, vol. 14, p. 39.
- Roy B. (1997). *L'aide à la décision aujourd'hui: que devrait-on en attendre?* Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision, Université Paris-Dauphine. Consulté sur <https://books.google.tn/books?id=9hosPwAACAAJ>
- Shani G., Gunawardana A. (2011). Evaluating recommendation systems. In *Recommender systems handbook*, p. 257–297. Springer.
- Shani G., Rokach L., Meisles A., Naamani L., Piratla N., Ben-Shimon D. (2007). Establishing user profiles in the mediascout recommender system. In *Computational intelligence and data mining, 2007. cidm 2007. IEEE symposium on*, p. 470–476.
- Sridevi K., Umarani R. (2013). Web personalization approaches: a survey. *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, n° 3.
- Symeonidis P., Malakoudis D. (2016). Moorec. com: Massive open online courses recommender system.
- Welch D. E., Welch L. S. (2008). The importance of language in international knowledge transfer. *Management International Review*, vol. 48, n° 3, p. 339–360.
- Yang D., Piergallini M., Howley I., Rose C. (2014). Forum thread recommendation for massive open online courses. In *Educational data mining 2014*, p. 257–260.

