



République Algérienne Démocratique et Populaire
Université Abou Bakr Belkaid– Tlemcen
Faculté des Sciences
Département d'Informatique

Mémoire de fin d'études

pour l'obtention du diplôme de Master en Informatique

Option : Système d'Information et de Connaissances (S.I.C)

Thème

Filtrage collaboratif des objets pédagogiques

Réalisé par :

- M^{elle} Zouag Rafika
- M^{elle} Khial Nabila

Présenté le 19 Septembre 2013 devant le jury composé de MM.

- Mr. Maatallah Hocine (Président)
- Mr. Boudefla Amine (Encadreur)
- Mr. Smahi Mohamed (Examineur)
- Mme Yebdri Zeyneb (Examineur)

Année universitaire: 2012-2013



Remerciement

Nous remercions « Allah » de nous avoir accordé des connaissances de la science et de nous avoir aidé à réaliser ce travail et qui, sans sa miséricorde, ce travail n'aurait pas abouti.

Au terme de ce modeste travail nous tenons à remercier chaleureusement et respectivement notre encadreur Mr. Boudefla Amine, qui en tant qu'encadreur de mémoire, s'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'il a bien voulu nous consacrer.

Nous adressons nos remerciements les plus sincères aux personnes qui nous ont apporté leur aide et qui ont contribué à l'élaboration de ce mémoire ainsi qu'à la réussite de cette formidable année universitaire.

Nous exprimons notre gratitude à tous les consultants et internautes rencontrés lors des recherches effectuées qui ont accepté de répondre à nos questions avec gentillesse.

Nous n'oublions pas nos parents pour leur contribution, leur soutien et leur patience.

Enfin, nous adressons nos plus sincères remerciements à tous nos proches et amis, qui nous ont toujours soutenue et encouragée au cours de la réalisation de ce mémoire.

Merci à tous et à toutes et que toute personne qui a contribué de près ou de loin à la réalisation de notre projet, trouve ici l'expression de nos sincères sentiments.

Rafika & Nabila





Dédicace

Je dédie ce mémoire ... 

À ma très chère mère : Zouag Rachida

Autant de phrases aussi expressives soient-elles ne sauraient montrer le degré d'amour et d'affection que j'éprouve pour toi. Tu m'as comblé avec ta tendresse et affection tout au long de mon parcours. Tu n'as cessé de me soutenir et de m'encourager durant toutes les années de mes études, tu as toujours été présente à mes cotés pour me consoler quand il fallait. Puisse Dieu le tout puissant te donner santé, bonheur et longue vie afin que je puisse te combler à mon tour.

À mon très cher père : Zouag Mohammed

Aucune dédicace ne saurait exprimer mon respect, mon amour éternel et ma considération pour les sacrifices que vous avez consenti pour mon instruction et mon bien être. JE vous remercie pour tout le soutien et l'amour que vous me porte depuis mon enfance et j'espère que votre bénédiction m'accompagne toujours. Puisse Dieu, le tout puissant, te préserver et t'accorder santé, longue vie et bonheur.

À mes chers et adorables frères Bilal et Soufiane, et ma chère sœur : Fouzia

De ma profonde tendresse et reconnaissance, je vous souhaite une



vie pleine de bonheur et de succès et que Dieu, le tout puissant, vous protège et vous garde.

À ma très chère amie : Dardek Amina

Un remerciement particulier et sincère pour Mon amie fidèle, qui m'a assisté dans les moments difficiles et m'a pris doucement par la main pour traverser ensemble des épreuves pénibles. Je te suis très reconnaissante, et je ne te remercierai jamais assez pour ton amabilité, ta générosité, ton aide précieuse, Je te souhaite un avenir plein de joie, de bonheur, de réussite et de sérénité.

À ma très chère amie : Slimani Meryem

En souvenir de notre sincère et profonde amitié et des moments agréables que nous avons passés ensemble. Veuillez trouver dans ce travail l'expression de mon respect le plus profond et mon affection la plus sincère.

À mon très chère binôme Khial Nabila

Qui avec elle j'ai partagé ce travail et j'ai trouvé la signification de l'amitié

À ma grande famille :

Mes tantes, mon oncle ainsi que mes cousins et cousines.

À la promotion d'informatique 2013 de l'université de Tlemcen

je cite en particulier :Imen Bensaid et Fatima Zohra Bendjelloul ;je vous souhaite une vie pleine de réussite, santé et de bonheur.

À tous mes enseignants dont la liste s'allonge d'une année à un autre

À tous les internes et résidents du Soufi Lamnawar, Tlemcen

À toutes les personnes qui ont participé à l'élaboration de ce travail à tous ceux que j'ai omis de citer

Zouag Rafika





Dédicace

Je dédie cet humble travail... 

A mes très chers parents pour leur amour et sacrifices,

A mon adorable frère Mohammed,

A ma très chère amie Chaimae,

A toute ma grande famille pour leur encouragement,

A mes proches amis Wafae, Khadidja, Aicha, Amina pour leurs patients,

A mon binôme Rafika,

A mes enseignants,

A toutes les personnes qui me connaissent.

Khial Nabila



Table des matières

Table des matières	1
Liste des figures	3
Liste des tableaux.....	3
Liste des abréviations.....	4
Introduction générale	5
Chapitre I :Généralités	7
1 Introduction	8
2 Systèmes de filtrage d'informations.....	8
3 Techniques de filtrage	9
3.1 Filtrage à base contenu.....	9
3.2 Filtrage collaboratif.....	10
3.2.1 Techniques de filtrage collaboratif	11
3.3 Filtrage hybride	14
4 Objets pédagogique	14
5 Profil d'utilisateur.....	15
5.1 Les caractéristiques	15
5.1.1 Caractéristiques Explicites.....	15
5.1.2 Caractéristiques implicites.....	15
5.2 Standards et modèles de données du profil existants.....	15
5.2.1 VCard.....	15
5.2.2 FOAF (Friend Of A Friend).....	16
5.2.3 PAPI Learner (Public and private information).....	16
6 Conclusion.....	17
Chapitre II : Etat de l'art	18
1 Introduction	19
2 Systèmes de filtrage collaboratif existant.....	19
2.1 Tapestry.....	19
2.2 GroupLens.....	19
2.3 Amazon	19
3 Le filtrage collaboratif basé sur le modèle des espaces de communautés.....	20
4 Le Filtrage Collaboratif distribué	20
5 Modèle multi-agents pour le filtrage collaboratif de l'information	22
6 Filtrage collaboratif à base du Framework de similarités sémantiques.....	22

7	Systèmes d'annotations collaboratives.....	23
8	Conclusion.....	24
	Chapitre III :Contribution Et Implémentation	25
1	Introduction	26
2	Contribution.....	26
3	Les outils de réalisation	27
3.1	Choix de langage de programmation Java	27
3.2	Choix du Framework Struts 2	28
3.3	Choix d'EasyPHP :	30
4	Architecture globale de système.....	31
5	Réalisation de l'application	31
5.1	Définition de E-Learning	31
5.2	La page d'accueil	31
5.3	Page d'administrateur.....	32
5.4	Connexion et inscription	33
5.5	Affichage et évaluation	33
5.6	Description de système de filtrage collaboratif.....	34
5.6.1	Calcul de distances.....	35
5.6.2	Calcul de prédiction	35
5.6.3	Exemple de recommandation.....	36
6	Conclusion.....	38
	Conclusion générale.....	39
	Perspectives :	39
	Les références	41
	ANNEXE A.....	44

Liste des figures

Figure I. 1: Schéma général du filtrage d'information.	9
Figure I. 2: Eléments de PAPI.	16
Figure II. 1: Architecture de l'algorithme RSB.....	21
Figure II. 2: Système avec topologie de voisinage hiérarchique.....	21
Figure II. 3: Similarité entre utilisateurs basée sur la composition de propriétés.....	23
Figure III. 1: L'architecture de Struts 2.....	29
Figure III. 2: La table « cours » de la base de données.....	30
Figure III. 3: La table « étudiant » de la base de données.....	30
Figure III. 4: La table « notes » de la base de données.....	31
Figure III. 5: Architecture générale d'un système de filtrage collaboratif.....	31
Figure III. 6: Page d'accueil de la plateforme E-Learning.....	32
Figure III. 7: La page d'administrateur.....	32
Figure III. 8: Connexion et inscription sur la plateforme.....	33
Figure III. 9: L'affichage et l'évaluation d'un cours.....	34
Figure III. 10: Recommandation des cours par le système.....	34
Figure A. 1: Le code de corrélation de pearson.....	45
Figure A. 2: Le code de prédiction.....	46

Liste des tableaux

Tableau I. 1: La matrice Utilisateur/Document.....	10
Tableau III. 1: Matrice des votes.....	35
Tableau III. 3: matrice cours-apprenant.....	36

Liste des abréviations

COCofil : Filtrage Collaboratif orientée vers la Communauté

FBC : Filtrage Basé sur le Contenu

FC: Filtrage Collaborative

FOAF: Friend of A Friend

GAMuS: Genetic auto-Adaptative Multi-agent System

GLRP: Group Lens Research Project

OP: Objets Pédagogique

PAPI Learner: Public and Private Information Learner

P2P: Peer-to-Peer

RDF: Resource Description Framework

RSB: RecTree SEM Bayesian

SMA: Système Multi-Agents

TA: Task Agents

VCard: VisitCard

XML: eXtensible Markup Language

Introduction générale

Introduction générale

1. Préambule

Les systèmes de recommandation sont nés de la volonté de pallier au problème de surcharge d'information du web. Ils ont pour but de préconiser à un utilisateur des ressources en lien avec ses goûts et ses attentes. L'objectif est à la fois de minimiser son temps passé à la recherche, mais aussi de lui suggérer des ressources pertinentes qu'il n'aurait pas spontanément consultées et ainsi accroître sa satisfaction globale.

Pour déterminer les ressources à recommander, plusieurs approches sont possibles : l'approche par contenu qui effectue des recommandations en comparant le contenu sémantique des ressources avec les goûts exprimés par l'utilisateur, et l'approche par filtrage collaboratif développé par le laboratoire Media Lab du MIT, cette approche effectue des recommandations par analyse à la fois des opinions de l'utilisateur sur les ressources qu'il a consultées ainsi que celles des autres utilisateurs sur les ressources qu'ils ont consultées. L'intérêt porté à cette approche s'est largement accru ces dernières années, dans ce mémoire nous nous intéressons à l'approche par filtrage collaboratif (FC).

2. Problématique

Comme nous l'avons indiqué précédemment, les systèmes de recommandation visent à personnaliser l'accès à l'information fournie par un système informatique. Pour atteindre cet objectif, les systèmes de recommandation peuvent notamment exploiter la technique du FC afin de modéliser les apprenants et leur recommander des ressources pertinentes en se basant sur les opinions de leurs voisins.

La problématique scientifique que nous traitons est liée à la recommandation des ressources à l'apprenant actif exploitant le filtrage collaboratif en se basant sur l'évaluation des autres apprenants et sur le calcul de distance de similarités avec tous les utilisateurs de système.

3. Objectifs

L'objectif principale de notre travail consiste à réaliser un système de recommandation collaborative des objets pédagogiques en se basant sur la corrélation de pearson l'un des algorithmes basés mémoire elle est la plus utilisé et la plus populaire car elle possède le

Introduction générale

meilleur comportement pour le filtrage collaboratif et sa popularité est liée à sa contribution à la performance des systèmes de recommandation.

Organisation

Ce mémoire contient trois chapitres qui se répartissent comme suit :

Dans le premier chapitre, un aperçu global sur le filtrage d'information et ses techniques en particulier le filtrage collaboratif, nous présentons ensuite, la description du profil utilisateur et de l'objet pédagogique.

Dans le deuxième chapitre nous présentons l'état de l'art sur les travaux relatifs au filtrage collaboratif.

Dans le troisième chapitre, nous commençons par notre contribution liée au le filtrage collaboratif des objets pédagogique. Ensuite, nous présentons l'architecture globale du notre système, enfin nous présentons une description sur la réalisation de notre application et nous détaillons l'algorithme de pearson permettant le calcul des similarités et des prédictions afin de recommander les ressources pertinentes à l'utilisateur actif.

Chapitre I :

Généralités

1 Introduction

Les systèmes de filtrage d'information ont connu une avancée significative ces dix dernières années, depuis les premiers systèmes collaboratifs classiques à ceux à base de contenu ou hybrides. Ils ont été largement investis dans divers domaines tels que le commerce électronique (livres, cinéma, musique, voyages, restauration, etc.).

Dans ce chapitre nous commençons par présenter un aperçu global du filtrage d'information et ses techniques en particulier le filtrage collaboratif, nous présentons ensuite, la description d'un objet pédagogique et d'un profil utilisateur.

2 Systèmes de filtrage d'informations

Les systèmes de filtrage d'information, ou système de recommandation (Recommender System), ont pour objectif de filtrer un flux entrant d'informations c'est-à-dire collecte, sélectionne, classe et suggère à chaque utilisateur les informations qui répondent à ses besoins de façon personnalisée [1].

Le filtrage d'information est l'expression utilisée pour décrire une variété de processus ayant pour but de fournir des informations à des personnes, répondant à leurs centres d'intérêt [2].

Le filtrage peut être vu comme la sélection d'informations pertinentes sur un flux entrant. Le système fait une « prédiction » quant à l'intérêt que présente l'information pour l'utilisateur. Cette prédiction s'appuie sur le « profil » de cet utilisateur et aboutit à une prise de décision : « recommander » ou « ne pas recommander » l'information. [3]

Les propriétés principales d'un système de filtrages sont :

- Filtrer l'information non pertinente et garder seulement ce qui est intéressant.
- Le flux d'information est dynamique (envoi d'alertes, ...).
- Personnalisation via la constitution et la maintenance d'un profil utilisateur, et prise en charge des intérêts à long terme.
- L'adéquation entre besoin et ressource peut être faite avec plusieurs dispositifs : indexation, contenu sémantique, évaluation, ...
- Interaction de l'utilisateur, avec l'évaluation des ressources recommandées, pour la mise à jour de son profil. [4]

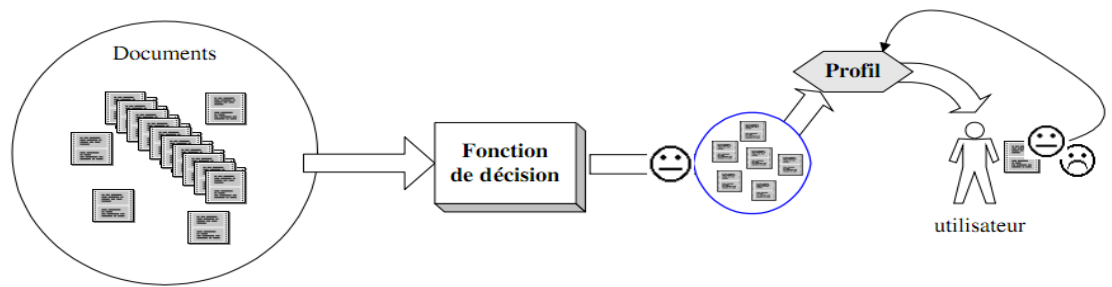


Figure I. 1: Schéma général du filtrage d'information. [1].

Aujourd'hui, ils existent de nombreux systèmes de filtrage appliqués existents dans différents domaines tel que la recherche documentaire, le commerce électronique. On peut citer à titre d'exemple quelques sites web populaires comme Amazon¹, Movielen², CiteSeer³, etc.

3 Techniques de filtrage

Le filtrage d'information est catégorisé en trois grandes approches :

- Filtrage à base de contenu.
- Filtrage collaboratif.
- Filtrage hybride.

Dans la suite, nous présentons plus en détails ces approches de filtrage, en particulier le filtrage collaboratif qui est au centre de notre travail.

3.1 Filtrage à base contenu

Les systèmes de filtrage à base de contenu recommandent des documents similaires à ceux que l'utilisateur a déjà appréciés. Ceci est calculé en comparant les centres d'intérêt des utilisateurs (exprimés implicitement à travers la surveillance de son comportement ou explicitement par exemple à travers un questionnaire) avec les métadonnées ou les caractéristiques utilisées pour représenter les ressources ou les produits, sans prendre en compte les avis et les informations concernant d'autres utilisateurs [5].

¹ <http://www.amazon.com>

² <http://citeseer.ist.psu.edu>

³ <http://movielens.umn.edu>

Deux fonctionnalités centrales ressortent de ce type de systèmes :

- La première étant la sélection de ressources correspondant au profil de l'utilisateur,
- La deuxième étant la mise à jour du profil de l'utilisateur après retour de pertinence des résultats [3].

Cependant, ce type de systèmes présente certaines limitations :

- L'effet « entonnoir » : La spécification des besoins de l'utilisateur empêche la diversité des sujets.
- Le filtrage des documents basé sur le contenu ne permet pas d'intégrer d'autres facteurs de pertinence (la qualité scientifique, le public visé, l'intérêt porté par l'utilisateur) que le facteur thématique.
- la difficulté d'indexation de documents multimédia cause une difficulté de recommander ce type de document.
- Problème de démarrage à froid : Un nouvel utilisateur du système éprouve des difficultés à exprimer son profil en spécifiant des thèmes qui l'intéressent [4].

3.2 Filtrage collaboratif

A l'opposé du filtrage basé sur le contenu, le filtrage collaboratif (Collaborative Filtering CF) prend en compte les « évaluations » que les utilisateurs ont données à certains des documents, pour recommander ces mêmes documents à d'autres utilisateurs. Ainsi, la fonction de prédiction F utilise la matrice des votes (utilisateur \times document) $U \times I \rightarrow [1,10]$ [6].

	S1	S2	S3
C1		6	7
C2	5	4	
C3	8		3
C4	5		
C5		2	

Tableau I. 1:La matrice Utilisateur/Document [3].

Un système de filtrage collaboratif est organisé comme suit :

- Collecter les appréciations des utilisateurs en général les évaluations est exprimée sous forme de notes,
- Intégrer ces informations dans les profils d'utilisateur,
- Et utiliser ceux-ci ensuite pour aider les utilisateurs dans leurs prochaines recherches [7].

3.2.1 Techniques de filtrage collaboratif

Breese et al. proposent une classification intéressante des techniques de filtrage collaboratif : les algorithmes basés « mémoire », et les algorithmes basés « modèle » [8].

a. Algorithmes basés « mémoire »

Les algorithmes basés mémoire utilisent l'ensemble de la base de données des évaluations des utilisateurs pour faire les prédictions : les évaluations de l'utilisateur actif sont prédits à partir d'informations partielles concernant l'utilisateur actif, et un ensemble de poids calculés à partir de la base de données des évaluations des utilisateurs. [8]

La prédiction de l'évaluation d'une ressource par un utilisateur sera calculée par la formule :

$$p_{a,j} = \bar{v}_a + \kappa \sum_{i=1}^n w(a,i)(v_{i,j} - \bar{v}_i)$$

- \bar{v}_a est l'évaluation moyenne de l'utilisateur 'a'
- $w(a, i)$ est un coefficient de corrélation entre l'utilisateur actif 'a' et l'utilisateur 'i' Coefficient de Pearson, Cosinus entre les vecteurs, les Moindres Carrés, ...
- \bar{v}_i est l'évaluation moyenne de l'utilisateur 'i'
- κ est un coefficient de normalisation, tel que la somme des poids est égal à 1.
- N est le nombre d'utilisateurs considérés pour le calcul

Le poids $w(a, i)$ est déterminé de façon variable, selon l'algorithme :

- ✓ l'algorithme basé sur la corrélation.
- ✓ l'algorithme basé sur la similarité de vecteurs.
 - i. **Corrélation de Pearson :** Le coefficient de Pearson a été utilisé dans le contexte des systèmes de filtrage collaboratif pour la première fois dans les recherches du GroupLens Research Project.

Le coefficient est calculé avec la formule suivante :

$$w(a, i) = \frac{\sum_j (v_{a,j} - \bar{v}_a)(v_{i,j} - \bar{v}_i)}{\sqrt{\sum_j (v_{a,j} - \bar{v}_a)^2 \sum_j (v_{i,j} - \bar{v}_i)^2}}$$

- $V_{a,i}$ est l'évaluation de l'utilisateur 'a' pour la ressource 'j'.
- \bar{V}_a est l'évaluation moyenne de l'utilisateur 'a'.

ii. Cosinus des vecteurs

Pour l'algorithme basé sur la similarité des vecteurs, le poids est calculé comme un cosinus entre les vecteurs formés par les évaluations des utilisateurs, comme suit [8] :

$$w(a, i) = \sum_j \frac{V_{a,j}}{\sqrt{\sum_{k \in I_a} V_{a,k}^2}} \frac{V_{i,j}}{\sqrt{\sum_{k \in I_i} V_{i,k}^2}}$$

- $V_{a,j}$ est l'évaluation de l'utilisateur a pour la ressource j.

iii. La distance de Spearman (utilisateurs)

Elle consiste à trouver un coefficient de corrélation, non pas entre les valeurs prises par les deux variables mais entre les rangs de ces valeurs. L'interprétation est identique à celle de la corrélation de Pearson [9].

b. Algorithmes basés modèle

Les algorithmes basés « modèle » utilisent la base de données des évaluations des utilisateurs pour estimer ou apprendre un modèle qui est alors utilisé pour les prédictions.

Du point de vue probabiliste, la tâche de prédiction d'une évaluation peut être vue comme le calcul de la valeur espérée d'une évaluation, étant donné ce que l'on sait d'un utilisateur. [8]

Il existe deux modèles probabilistes : le modèle à base de clusters, et le modèle à base de réseau bayésien :

i. Clustering

L'idée du modèle cluster est de regrouper (en clusters) les gens ayant les mêmes goûts, et de regrouper (en clusters) les articles portant sur les mêmes sujets, ou qui tendent à plaire aux mêmes personnes. Ainsi, pour prédire la note qu'un utilisateur donnera à un article, on pourra utiliser les avis des utilisateurs qui appartiennent à son groupe [7].

➤ **K-Moyennes**

L'idée principale de l'algorithme des K-moyennes est de classer des objets en K classes. Cet algorithme commence par choisir au hasard K centres de gravité. Puis, construire les classes initiales autour de ces centres : chaque objet appartient à la classe dont le centre est le plus proche. A chaque itération, il faut recalculer les centres en fonction de la variance intra-classe, et former les nouvelles classes jusqu'à l'obtention plus de changement de partition [1].

ii. Réseaux de Bayes

Il est possible d'utiliser les réseaux de Bayes et les arbres de décision dans le contexte du filtrage collaboratif [10]. D'après Heckerman Un réseau de Bayes est un graphe acyclique dirigé qui représente une distribution de probabilités de dépendance entre un ensemble de variables. Chaque nœud dans le graphe représente une variable, et chaque arc une dépendance directe entre variables. Ainsi, chaque variable est indépendante de ses non-descendants dans le graphe, étant donné l'état de ses parents [7].

c. Algorithmes hybrides

Les approches hybrides viennent combiner efficacement les deux approches pour limiter les inconvénients.

i. Approche graphique : Horting

Le filtrage collaboratif par **Horting** est une approche basée sur un graphe de relations de similarité (représentées par les arcs) entre les utilisateurs (correspondant aux nœuds du graphe). La notion d'influence (arcs) se décline sous forme de deux contraintes :

- la contrainte de Horting : ne considérer que les utilisateurs ayant un grand nombre d'évaluations communes.
- la contrainte de prédictibilité ajoute à la notion de « Horting » une information sur le degré de ressemblance entre deux utilisateurs en se basant sur la distance de Manhattan. [12]

ii. Diagnostique de personnalité

Pennock et al. ont introduit une méthode de modélisation de la personnalité permettant de considérer les évaluations des utilisateurs avec un bruit Gaussien. Ainsi, il est possible de prendre en compte des paramètres extérieurs tels que l'humeur de l'utilisateur [13]. Par exemple les évaluations sont considérées comme des symptômes et le type de personnalité comme une maladie. De cette manière, déterminer cette personnalité revient à identifier la cause la plus probable des évaluations. Le principal bénéfice de cette approche réside dans le fait que les prédictions du modèle sont exprimées explicitement ce qui rend possible leur modification et leur validation [12].

3.3 Filtrage hybride

L'approche de filtrage hybride repose sur l'idée de tirer profit des avantages des deux approches précédentes, en résolvant les problèmes qui leur sont liés. En fait, ces deux approches paraissent complémentaires. Les chercheurs du domaine estiment que le fait de combiner les deux méthodes pourrait être très bénéfique. D'où l'émergence de plusieurs techniques d'hybridation dont l'objectif consiste à combiner les deux approches (filtrage collaboratif et filtrage basé sur le contenu) de manière efficace [14].

4 Objets pédagogique

En E-Learning, la forme fondamentale d'une ressource pédagogique est l'« objet pédagogique » (OP); Le groupe de travail des IEEE-LTSC a proposé une définition d'un OP qui est la suivante : « Un objet pédagogique est défini comme toute entité numérique ou non qui peut être utilisée, réutilisée ou référencée pendant des activités d'apprentissage assistées par ordinateur (enseignement intelligent assisté par ordinateur, environnements d'enseignement interactifs, systèmes d'enseignement à distance, environnements d'apprentissage collaboratif) » [15].

5 Profil d'utilisateur

Un profil utilisateur (ou encore un modèle utilisateur) est une collection de données personnelles associées à un utilisateur spécifique. Un profil se réfère donc à la représentation numérique explicite de l'identité d'une personne. Un modèle utilisateur est une source de connaissance qui contient des acquisitions sur tous les aspects de l'utilisateur qui peuvent être utiles pour le comportement du système [16].

5.1 Les caractéristiques

Les caractéristiques d'un profil utilisateur peuvent être décomposées en deux groupes :

5.1.1 Caractéristiques Explicites

Ces traits sont représentés par : les acquis de l'utilisateur, ses connaissances, ses buts, ses préférences. Auxquels on peut ajouter les critères de choix et de centre d'intérêt. Les caractéristiques explicites sont en général objectives [4].

5.1.2 Caractéristiques implicites

Ce sont les caractéristiques qui sont induites à partir de l'activité de navigation, d'évaluation, marquage avec des tags, réponse à des questions, etc. L'avantage d'utiliser des techniques implicites se fait sentir car l'utilisateur est déchargé de certaines actions, telles que la définition de ces préférences, de son caractère, etc., Les caractéristiques implicites sont en général subjectives [4].

5.2 Standards et modèles de données du profil existants

Dans la suite, nous allons présenter certains standards existant de modèle utilisateur existants avec leurs caractéristiques

5.2.1 VCard

VCard (VisitCard) est un format ouvert pour l'échange des données personnelles, il est utilisé dans des applications telles que la messagerie Internet, les navigateurs Web, les appareils mobiles, les logiciels de carnet d'adresses (Outlook, Thunderbird, ...).

Avec VCard, il est possible de représenter des informations personnelles (nom, prénom, date de naissance, adresse, e-mail, téléphone, ...), des éléments tels que les images, les logos d'entreprise, ainsi de suite [17].

Il est possible de décrire les données VCard dans un format RDF ou XML [18].

5.2.2 FOAF (Friend Of A Friend)

FOAF ou «ami d'un ami» est un vocabulaire basé sur RDF, défini dans le cadre d'un projet open source, permettant de décrire des personnes et les relations qu'elles entretiennent entre elles. Il a été développé pour la construction de groupes sociaux [19].

FOAF distingue 5 catégories pour décrire un profil [16]:

- FOAF Basics de base comprend une description de base telle que nom, adresse e-mail, images.
- Personal Information décrit plus d'informations personnelles telles que le blog, les intérêts, les publications et les relations aux autres profils qui connaissent cette personne.
- Online Accounts décrit les informations sur les comptes qu'une personne possède.
- Projets and Groups définit les informations sur les projets, les groupes ou les organisations dans lesquelles la personne est membre.
- Documents and Images décrit les documents et les images relatives à l'apprenant, par exemple: document de profil, logo...

5.2.3 PAPI Learner (Public and private information)

PAPI Learner est un standard proposé par le groupe Learner Model Working Group de l'IEEE, qui décrit les informations sur l'apprenant utiles pour la communication entre les systèmes coopératifs. [16]

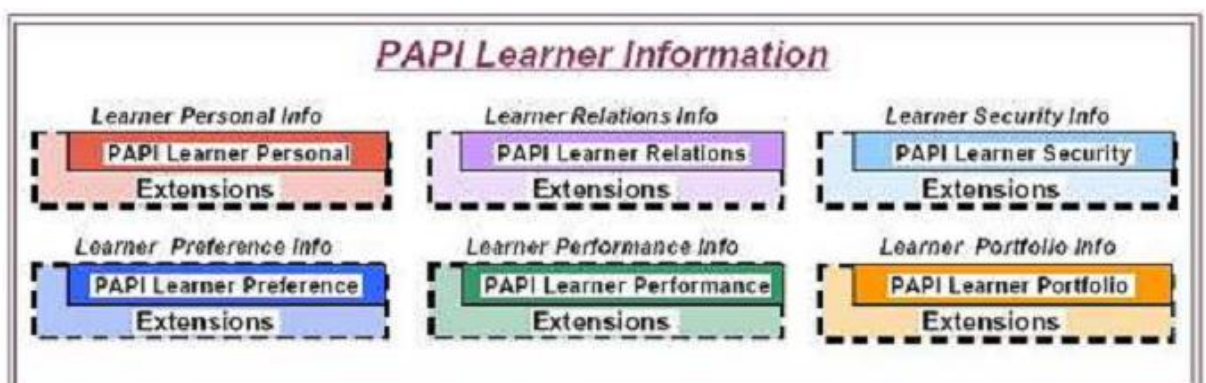


Figure I. 2:Eléments de PAPI [16].

PAPI décompose le profil en six catégories :

- Informations personnelles contient des informations générales sur l'étudiant, par exemple nom, adresse, etc.
- Relations contient des relations apprenant avec d'autres personnes, par exemple camarade, professeur, etc.
- Sécurité occupe des fonctions de sécurité des apprenants et des droits d'accès, par exemple clés publiques et privées.
- Préférences occupe l'information publique sur les préférences de l'apprenant, par exemple les styles d'apprentissage ou la langue.
- Performances détient le record de la performance de l'apprenant, qui peut être utilisée pour une évaluation ou d'identification des expériences, par exemple les notes d'apprentissage, les rapports, ou la certification, etc.
- Portfolio décrit les travaux et les projets de l'apprenant. Il a pour but d'accéder à son histoire et son expérience précédente. [16]

6 Conclusion

Dans les dernières années Les systèmes de filtrage d'information ont connu une avancée notable Qu'ils soient basés sur le contenu ou bien collaboratifs.

Le filtrage collaboratif est sans doute le type le plus utilisé il peut s'avérer très utile pour conseiller les utilisateurs sur certains textes qu'ils n'ont pas lus, et dont on sait qu'ils, intéressent son groupe. Déterminer un profil thématique pour chaque utilisateur permet par contre d'évaluer si un texte est pertinent ou non pour cet utilisateur, en fonction de ses centres d'intérêts identifiés. Or, ces deux aspects du profil paraissent importants. Connaître les goûts thématiques personnels paraît la meilleure méthode pour fournir une aide spécialisée au client, mais associer les clients entre eux permet de faire bénéficier chacun des opinions des autres.

Chapitre II :

Etat de l'art

1 Introduction

Les systèmes de recommandation ont une importance majeure pour le web d'aujourd'hui car ils permettent de cibler les préférences de l'utilisateur et de filtrer l'information qui lui parvient. C'est pour cela que plusieurs travaux ont été menés et le filtrage d'information en particulier le filtrage collaboratif est devenu un axe de recherche très actif.

Ce chapitre a pour objectif de faire un tour d'horizon, non exhaustif, des systèmes de recommandations collaboratives liées au domaine de la recherche d'information, en évoquant leur origine et leurs applications.

2 Systèmes de filtrage collaboratif existant :

2.1 Tapestry

Le concept du filtrage collaboratif a été lancé avec le projet Tapestry à Xerox Parc. La gestion des e-mails est sa motivation première.

L'implication de l'utilisateur n'est pas limitée à fournir de simples jugements binaires d'acceptation ou de rejet. Il donne la possibilité de faire des annotations en texte libre ou des appréciations dans le style « J'ai bien aimé » ou « Je déteste », ainsi les utilisateurs peuvent transmettre des jugements sur la valeur des documents qu'ils lisent. Les autres utilisateurs peuvent alors opérer des recherches parmi ces documents non seulement sur la base de leur contenu, mais également sur la base des jugements qu'ont portés d'autres utilisateurs à leur sujet. [8]

2.2 GroupLens

GroupLens système expérimental de l'université du Minnesota. Il est semblable dans son esprit à Tapestry : les lecteurs sont appelés à noter les articles qu'ils lisent sur une échelle numérique de cinq niveaux. Le système trouve alors des corrélations entre les différents utilisateurs RI d'estimations et identifie des groupes d'utilisateurs dont les intérêts sont semblables, et ensuite il emploie ces estimations pour prédire l'intérêt que porteront les lecteurs à chaque article. [8]

2.3 Amazon

Amazon.com, Inc. est une entreprise de commerce électronique américaine basée à Seattle. Sa spécialité la plus connue est la vente de livres, mais elle est diversifiée dans

d'autres produits, notamment dans la vente de tous types de produits culturels : disques CD, musique en téléchargement, DVD, appareils photos numériques, informatique et dans l'équipement de la maison, etc. [20].

Le système d'Amazon était un système passif qui se basait sur les achats des gens pour construire la matrice de relation entre les objets. Le système de filtrage collaboratif a été popularisé par Amazon avec la fonctionnalité "les gens qui ont acheté x ont aussi acheté y". [21]

3 Le filtrage collaboratif basé sur le modèle des espaces de communautés

L'équipe **MRIM** à mener des travaux autour de la perception des communautés dans un système de filtrage collaboratif. Parmi ces travaux :

Denos et al. en 2004 réalisent un système de recommandation « COCoFil1 » basé sur le filtrage collaboratif et offre des nouvelles fonctionnalités orientées vers la notion de communauté comme Le carnet d'adresses et la perception des autres[22]. Ensuite An-Te Nguyen en 2006 a développé un nouveau système de filtrage collaboratif basé sur le modèle des espaces de communautés intitulé « COCoFil2 » qui permet de gérer efficacement les communautés et vérifier les trois caractéristiques suivantes : gestion des communautés explicites, formation multiple des communautés selon divers critères et efficacité du positionnement des utilisateurs au sein des communautés[1]. Au cours du projet **APMD** Goundo et al. en 2007 ont proposés de concevoir et développer une application « COCoFil3 », permet de définir et valider des modalités d'interaction entre un utilisateur et un moteur de recommandations fondé sur le principe de communautés d'utilisateurs multidimensionnelles.[23]

4 Le Filtrage Collaboratif distribué

Castagnos et al. en 2004 ont proposé l'algorithme RSB (RecTree SEM Bayesian) qui permet d'effectuer du filtrage collaboratif de façon distribué sur de grosses bases de données et dans un délai relativement court. Il importe de mesurer la satisfaction des utilisateurs par rapport aux ressources que le système leur propose. Le sigle RSB reflète donc l'association de l'algorithme de classification **RecTree** (scindé en deux parties, respectivement du côté serveur et du côté client), du procédé d'apprentissage **SEM** et d'une méthode de filtrage collaboratif à base de réseaux **Bayesiens**. [24]

L'architecture de l'algorithme est représentée sur la figure suivante :

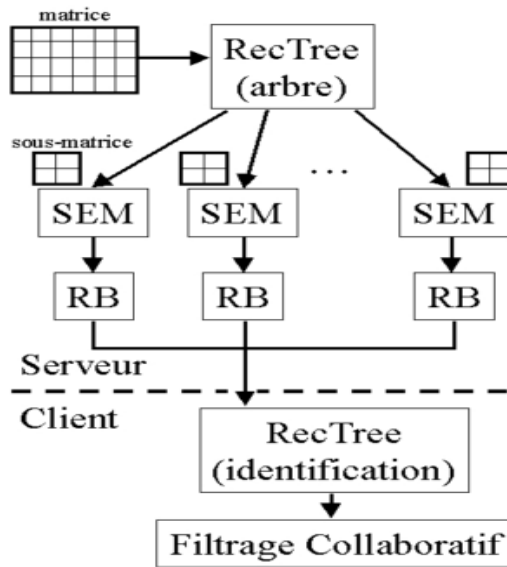


Figure II. 1:Architecture de l'algorithme RSB. [24]

Berkovsky et al. en 2006 proposent un algorithme de filtrage collaboratif distribué adapté aux architectures pair-à-pair et reposant sur une topologie de voisinage hiérarchique. [25]

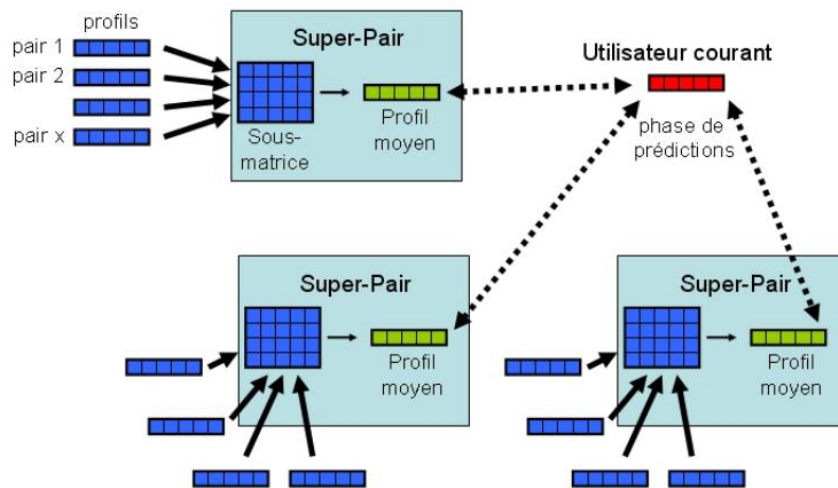


Figure II. 2:Système avec topologie de voisinage hiérarchique. [12]

Leur modèle repose sur une méthode d'agrégation des profils améliorant la confidentialité au sein de systèmes de recommandations pair-à-pair. Le système commence par élire des supers pairs dont le rôle est de calculer un profil moyen pour une sous-population. Dans ce modèle, seuls les super-pairs ont accès aux profils des

autres pairs. Il n'utilise que les profils moyens constitués, plutôt que des matrices regroupant l'ensemble des votes individuels. [25]

5 Modèle multi-agents pour le filtrage collaboratif de l'information

Zaier en 2003 propose le modèle cognitif qui est une plate-forme d'agents collaboratifs pour le filtrage hybride de l'information basée sur une architecture Peer-To-Peer. Il a considéré le système comme un ensemble d'agents indépendants, un agent pour chaque utilisateur du système. Il s'agit d'un système où l'utilisateur est passif et où les informations lui sont délivrées automatiquement (la technologie « Push » : comme les services d'alerte). [26]

Zaier en 2010 propose une nouvelle architecture pair-à-pair, de filtrage collaboratif, basée sur la discrimination du voisinage. Ainsi il a étudié l'évolution de la performance, de la couverture et de la qualité des prédictions pour différentes techniques de recommandation. En outre, il a identifié la méthode de recommandation la plus efficace pour cette nouvelle architecture pair-à-pair. [27]

Boussebough en 2011 a proposé un système qui est composé d'un ensemble de Task Agents et d'un agent médiateur selon le modèle GAMuS (Genetic auto-Adaptative Multi-agent System) une approche pour la construction de systèmes multi-agents auto-adaptatifs. Chaque TA (Task Agents) représente un utilisateur qui est défini par son profil et qui communique de manière directe avec d'autres utilisateurs ayant des centres d'intérêt proches des siens. Les utilisateurs doivent pouvoir s'échanger librement tout type de ressources documentaires stockées au niveau de leurs propres machines et effectuer des recommandations les uns aux autres. [28]

Mazyad en 2013 Propose une méthode de formation de groupe comme solution pour résoudre les problèmes de coordination d'agents dans le contexte d'une architecture P2P. Puis ils ont utilisé l'association de SMA et P2P pour concevoir un système d'apprentissage collaboratif à distance. [29]

6 Filtrage collaboratif à base du Framework de similarités sémantiques

Dans le but d'améliorer la qualité des systèmes de recommandation en termes de précision et de couverture Nouali et al en 2009 présentent une approche qui utilise

l'infrastructure «web sémantique». Ils ont proposé un Framework qui permet la composition de plusieurs mesures de similarité pour définir la similarité entre 2 entités (entre utilisateurs ou entre ressources). Comme exemple pour la définition de la similarité entre deux utilisateurs il est possible de composer la similarité entre plusieurs attributs pondérés : l'âge, la langue, la région, les préférences, etc. [4]

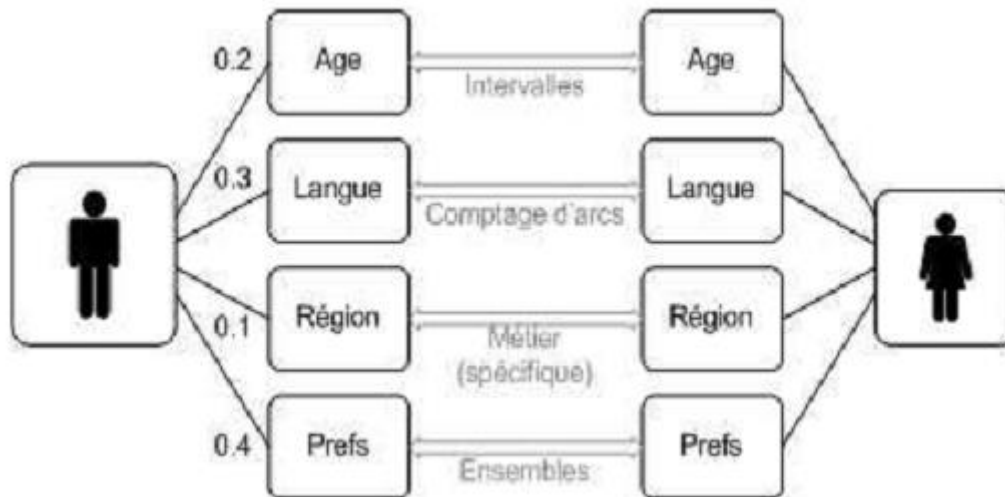


Figure II. 3: Similarité entre utilisateurs basée sur la composition de propriétés. [4]

7 Systèmes d'annotations collaboratives

Karen et al. en 2008 proposent d'étendre la matrice des votes avec les annotations comme troisième dimension. Cette matrice 3-dimensionnelle est ensuite décomposée en trois sous-matrices : (utilisateurs, documents), (utilisateurs, annotations) et (documents, annotations). [30]

Pour le calcul de la prédiction, les auteurs utilisent la combinaison de la similarité entre utilisateurs et celle entre documents mais cette fois sur la matrice étendue.

Szomszor et al. en 2007 proposent l'utilisation des annotations pour enrichir le profil utilisateur. Chaque utilisateur et même les documents disposent d'un nuage d'annotations. La recommandation est ainsi calculée en utilisant la similarité sémantique entre les deux nuages. [31]

Bouzidi et al en 2006 présentent un outil SMARTNotes qui supporte la collaboration entre les usagers via des forums de discussion à base d'annotation. Dans SMARTNotes pour associer une annotation à son contexte dans le document annoté, une base de liens a été définie associant le contenu de l'annotation qui est enregistré dans la base des

annotations (forum de discussion, question, réponse, etc.) à la partie du cours correspondante. Cette architecture permet d'afficher au lecteur du document annoté les annotations associées d'une manière transparente et sans aucune modification du document source [32].

8 Conclusion

D'autres systèmes d'importances variables sont disponibles et deviennent de plus en plus populaires (My Yahoo, Miningco.com, etc.). Ils offrent à leurs utilisateurs des possibilités de personnalisation d'interface, toutefois limitées, ou leur suggèrent suivant leurs centres d'intérêts constatés (explicitement ou implicitement) des ressources qui pourraient les intéresser.

Après avoir présenté les travaux existant sur le filtrage collaboratif, nous allons décrire dans le chapitre suivant l'approche générique que nous proposons, notre contexte applicatif proposés à travers ce travail.

Chapitre III :

Contribution

Et

Implémentation

1 Introduction

Les moteurs de recherche sur Internet ont permis la création et le partage d'informations. Leur manque de pertinence a donné naissance à de nouveaux outils comme le filtrage collaboratif, ce système était destiné à l'origine, à la recherche par un utilisateur de personnes partageant les mêmes centres d'intérêt, ou d'un expert sur un sujet donné.

Dans ce dernier chapitre, nous entamons effectivement notre contribution relative à la recommandation collaborative des objets pédagogique aux apprenants sur une plateforme e-Learning. Ensuite, nous détaillons l'algorithme utilisée pour le calcul des voisinages et des prédictions, les deux points essentiels dans un système de filtrage collaboratif, ainsi nous présentons l'architecture globale du notre système. Enfin, nous décrivons la réalisation de notre application.

2 Contribution

Nous nous intéressons à la recherche d'informations sur une plateforme e-Learning, pour objectif de déterminer quelles ressources sont pertinentes pour un utilisateur. Nos travaux s'inscrivent dans le domaine des systèmes de recommandation avec l'objectif de prédire la pertinence d'une ressource (Objet pédagogique) pour un utilisateur particulier à partir de l'observation de ses appréciations (les votes) qu'il dépose lors de ses consultations.

Parmi les techniques possibles, nous avons choisi d'utiliser le filtrage collaboratif, cette approche résout les problèmes de l'approche basée sur le contenu (difficulté d'indexation, incapacité à traiter d'autres critères, effet « entonnoir ») en se basant sur les évaluations des utilisateurs plutôt que sur le contenu, et donc de filtrer tout type d'information (ressources physiques, images, vidéos, etc.) ainsi l'effet entonnoir est atténué par le filtrage collaboratif, car tout document évalué par une personne peut être recommandé. L'utilisateur peut bénéficier de nouveaux axes de recherche auxquels il n'a jamais pensé, le filtrage collaboratif permet aussi d'exprimer d'autres facteurs et critères tels que : la qualité de l'information, le public visé, la zone géographique, etc. Chose qui n'est pas possible dans un système de filtrage thématique.

Le principe du filtrage collaboratif est d'implanter informatiquement le principe du bouche-à-oreille proposé par J. L. Herlocker et al en 2004. Il détermine pour l'utilisateur actif U_i quels sont les utilisateurs similaires.

Les utilisateurs de système fournissent des évaluations des documents, sous forme de notes, pour constituer leur profil. Ces estimations sont utilisés à la fois pour mesurer des similarités entre utilisateurs et pour prédire les votes que les utilisateurs donneraient aux ressources qu'ils n'ont pas vues. Ces votes estimés sont ensuite exploités pour effectuer des recommandations.

Le filtrage collaboratif emploie des méthodes pour faire des recommandations à un utilisateur, en se fondant sur la corrélation entre son propre profil personnel et les profils d'autres utilisateurs qui présentent des intérêts et goûts semblables.

Ces méthodes sont classées en deux familles principales : les algorithmes basés mémoire est les algorithmes basés modèle, dans le cadre de notre travail en particulier le filtrage collaboratif basé mémoire.

Plusieurs mesures ont été exploitées dans le cadre du FC basé sur la mémoire dans le but d'évaluer les similarités d'appréciations entre utilisateurs et identifier les utilisateurs voisins (les plus proches). Parmi ces mesures nous avons choisi la corrélation de Pearson.

3 Les outils de réalisation

Dans cette partie nous allons présenter les principaux outils utilisés pour la mise en place de notre application.

La réalisation de cette application a été faite sous la plateforme Java on se basant sur le Framework Struts2 avec l'utilisation d'EasyPHP comme serveur de base de données.

3.1 Choix de langage de programmation Java

Les modules conçus ont été réalisés sous Java dont les principales vertus, sont résumées dans les points suivants :

- Java est un langage orienté objet : un programme Java est centré complètement sur les objets et fournit un ensemble prédéfini de classes facilitant la manipulation des entrées-sorties, la programmation réseau, système, graphique. Excepté

les types de données fondamentaux, tout est objet ! Seule l'héritage simple existe en Java, pour utiliser l'héritage multiple, on utilisera les interfaces.

- Le langage Java est distribué : il est conçu pour développer des applications en réseau, les manipulations des objets distants ou locaux se font de la même manière.
- Le langage Java est robuste et sûr : Destiné pour des applications réseaux, la sécurité dans Java est un aspect primordial, il ne supprime pas tous les problèmes de sécurité mais les réduit fortement. [33]

Nous avons utilisé un éditeur de Java appelé **NetBeans version 6.9.1** un environnement de développement intégré (EDI), placé en *open source* par Sun en juin 2000 sous licence CDDL et GPLv2 (Common Development and Distribution License). En plus de Java, NetBeans permet également de supporter différents autres langages, comme Python, C, C++, JavaScript, XML, Ruby, PHP et HTML. Il comprend toutes les caractéristiques d'un IDE moderne (éditeur en couleur, projets multi-langage, refactoring, éditeur graphique d'interfaces et de pages Web). [34]

3.2 Choix du Framework Struts 2

Le Framework web Apache Struts est un logiciel gratuit open-source pour créer des applications web Java basées sur Java Server Pages (JSP).

Struts 2 implémente le modèle d'architecture dit MVC 2 (Modèle - Vue - Contrôleur), dans le but de séparer la partie Modèle (programmation, traitement des informations) de la partie Présentation (affichage). Le modèle contient les objets représentant les données, c'est-à-dire toutes les fonctionnalités de l'application. La vue représente l'état actuel du modèle, son aspect. Le contrôleur gère les interactions utilisateur/interface, crée et manipule le modèle et ses objets. [35]

La Figure III.1 schématise l'architecture de Struts 2 :

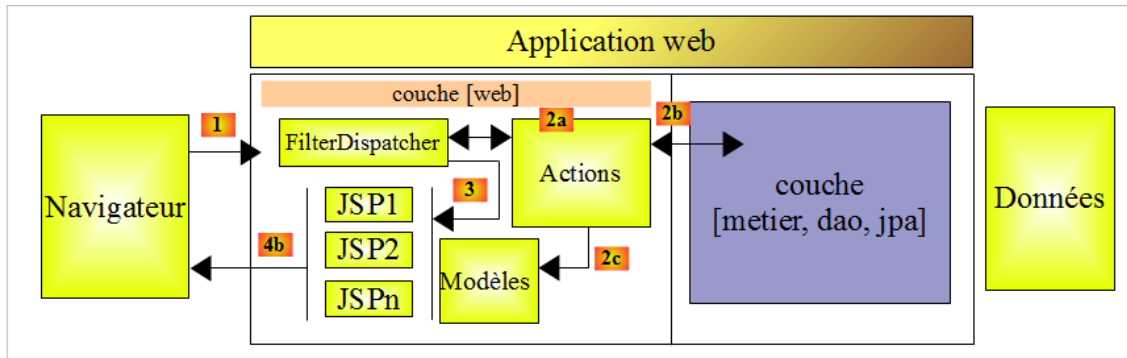


Figure III. 1:L'architecture de Struts 2 [36]

Le traitement d'une demande d'un client se déroule de la façon suivante :

Les Url demandées sont de la forme *http://machine:port/contexte/rep1/rep2/.../Action*. Le chemin *[/rep1/rep2/.../Action]* doit correspondre à une action définie dans un fichier de configuration de Struts 2, sinon elle est refusée. Une action est définie dans un fichier Xml sous la forme suivante :

1. `<package name="actions" namespace="/actions" extends="struts-default">`
2. `<action name="Action1" class="actions.Action1">`
3. `<result name="page1">/vues/Page1.jsp</result>`
4. `<result name="page2">/vues/Page2.jsp</result>`
5. `</action>`
6. `</package>`

Supposons qu'un Url soit demandée. Les étapes suivantes sont alors exécutées :

- **Demande :** le client navigateur fait une demande au contrôleur [FilterDispatcher]. Celui-ci voit passer toutes les demandes des clients. C'est la porte d'entrée de l'application. C'est le C de MVC.
- **Traitement :** le contrôleur C consulte son fichier de configuration et découvre que l'action *actions/Action1* existe, il instancie [2a] une classe de type [*actions.Action1*], si l'Url demandée est accompagnée alors le contrôleur C affecte ces paramètres à la classe, le contrôleur C demande à la méthode [*String execute()*] de la classe [*actions.Action1*] de s'exécuter, une fois la demande du client traitée, celle-ci peut appeler diverses réponses, la méthode *execute()* rend au contrôleur C un résultat de type chaîne de caractères appelée clé de naviga-

tion, la méthode *execute()* va également mettre à jour le modèle M [2c] que va exploiter la page JSP qui va être envoyée en réponse à l'utilisateur.

- **Réponse** : le contrôleur C demande à la page JSP correspondant à la clé de navigation de s'afficher. C'est la vue, le V de MVC. La page JSP utilise un modèle M pour initialiser les parties dynamiques de la réponse qu'elle doit envoyer au client. [36]

3.3 Choix d'EasyPHP

C'est un environnement comprenant un serveur Web (Apache), un serveur SQL (MySQL), un interpréteur de script PHP et un administrateur de base SQL (phpMyAdmin). Cet environnement permet de faire fonctionner en local un site Internet développé en PHP sans à avoir à se connecter à un serveur externe. Cela peut être utile dans le cadre du développement d'un site pour tester différentes versions ou pour rendre accessible un site Internet sur un réseau fermé, comme un Intranet. EasyPHP est gratuit et s'installe comme une application même si ce n'est pas une application. [37]

On a choisi EasyPHP comme outil pour créer la base de données qui constitue le langage intermédiaire entre cette base et l'utilisateur de la base.

Notre base de données contient les tables suivantes :

- **Cours** (Cours ID, CoursName, CoursImage, Lien, description, filière, module, prédiction).

		coursID	coursName	coursImage	lien	Description	filiere	module	pred
<input type="checkbox"/>		1	Relationnel	Rel.PNG	http://127.0.0.1:8080/PFE/result1.jsp	Module: Base de données / Réalisé par: Amel HA...	info	bdd	0
<input type="checkbox"/>		2	SQL	SQL.PNG	http://127.0.0.1:8080/PFE/result2.jsp	Module: Base de données / Réalisé par: Amel HA...	info	bdd	0
<input type="checkbox"/>		3	Trigger	Trigger.PNG	http://127.0.0.1:8080/PFE/result3.jsp	Module: Base de données / Réalisé par: Amel HA...	info	bdd	0

Figure III. 2: La table « cours » de la base de données

- **Étudiant** (Login, mail, password, ID, Nom, Prénom).

			login	mail	pwd	id	nom	prenom
<input type="checkbox"/>			admin	admin@live.fr	admin	12	admin	admin
<input type="checkbox"/>			amina	amina.dardek@hotmail.com	amina	22	dardek	amina
<input type="checkbox"/>			nabila	nabila@live.fr	nabila	4	khial	nabila
<input type="checkbox"/>			et2	rafika13420@live.fr	a	3	zouag	rafika

Figure III. 3: La table « étudiant » de la base de données.

- **News**(Actualités).
- **Notes** (etudiant, cours1, cours2, cours3, cours4, cours5, cours6, cours7, Login, Corrélation).

	etudiant	cour1	cour2	cour3	cour4	cour5	cour6	cour7	login	correlation
<input type="checkbox"/>	et1	7	0	8	7	0	0	0	et2	-0.821107731803577
<input type="checkbox"/>	et2	0	5	0	2	5	8	8	amina	1
<input type="checkbox"/>	et3	0	2	2	5	0	0	0	nabila	-0.8

Figure III. 4:La table « notes » de la base de données.

4 Architecture globale de système

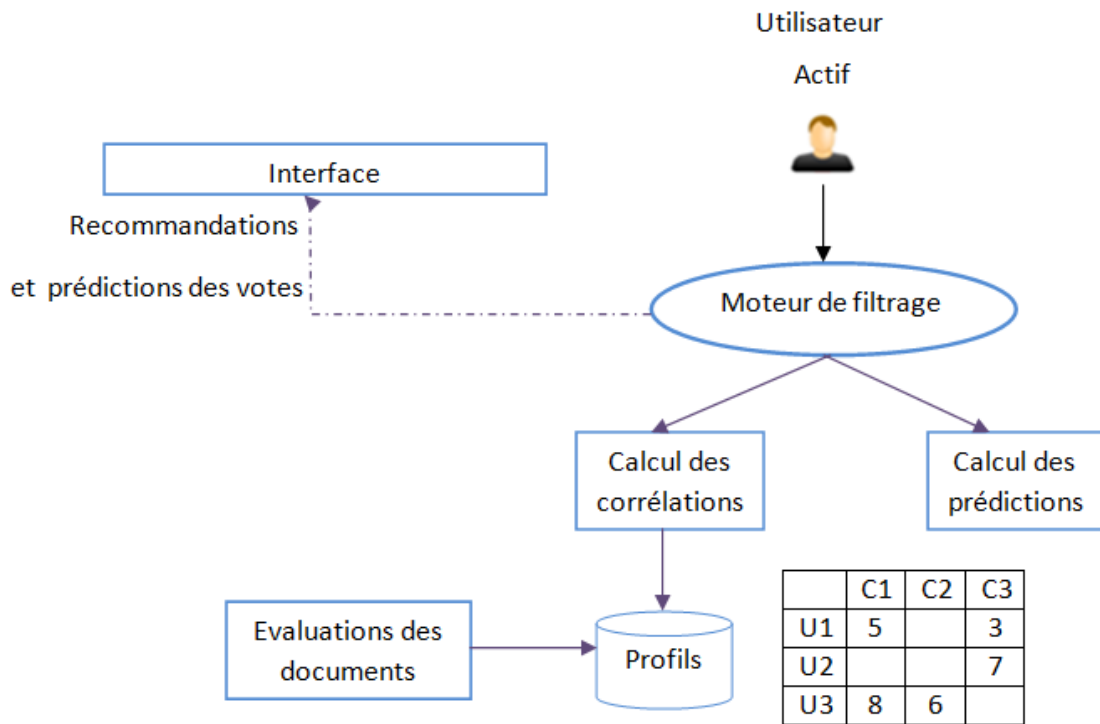


Figure III. 5:Architecture générale d'un système de filtrage collaboratif.

5 Réalisation de l'application

5.1 Définition de E-Learning

Le e-Learning c'est la possibilité de suivre un programme de formation à distance, en auto-formation ou accompagné, de manière individuelle ou collective. Le e-Learning s'appuie sur internet et les outils multimédia pour offrir des modules de formation courts, progressifs, adaptés aux niveaux et besoins des apprenants. [38]

5.2 La page d'accueil

La Figure III.6 présente la page d'accueil de la plateforme E-Learning :



Figure III. 6:Page d'accueil de la plateforme E-Learning.

Sur une plateforme E-Learning on peut lire des cours en ligne. Pour cela, il suffit le plus souvent de s'inscrire sur la plateforme puis de connecter avec un login et un mot de passe.

5.3 Page d'administrateur

Dans le partie administrateur, le nombre des utilisateurs est limitée seul les administrateurs auront le droit d'accéder à cette partie via un nom d'utilisateur et un mot de passe. Dans cette partie l'administrateur peut ajouter des cours à la base de données.

Insérer les informations concernant le cours

Nom du cours	Description	Images	Filiere	Module	Lien
Relationnel	Module: Base de données / Réalisé par: Amel KALFAOUI /	Rel.PNG	info	bdd	http://127.0.0.1:8080/PFE/result1.jpg
SQL	Module: Base de données / Réalisé par: Amel KALFAOUI /	SQL.PNG	info	bdd	http://127.0.0.1:8080/PFE/result2.jpg
Trigger	Module: Base de données / Réalisé par: Amel KALFAOUI /	Trigger.PNG	info	bdd	http://127.0.0.1:8080/PFE/result3.jpg
XQuery	Module: Base de données / Réalisé par: Amel KALFAOUI /	Xquery.PNG	info	bdd	http://127.0.0.1:8080/PFE/result4.jpg
Le	Module: Intelligence Artificiel / Réalisé par: Belabed Amine /	OWL.PNG	info	IA	http://127.0.0.1:8080/PFE/result5.jpg
Ontologies	Module: Intelligence Artificiel / Réalisé par: Belabed Amine /	ontologie.PNG	info	IA	http://127.0.0.1:8080/PFE/result6.jpg
RDF	Module: Intelligence Artificiel / Réalisé par: Belabed Amine /	RDF.PNG	info	IA	http://127.0.0.1:8080/PFE/result7.jpg
RDFS	Module: Intelligence Artificiel / Réalisé par: Belabed Amine /	RDFS.PNG	info	IA	http://127.0.0.1:8080/PFE/result8.jpg
XML	Module: Intelligence Artificiel / Réalisé par: Belabed Amine /	XML.PNG	info	IA	http://127.0.0.1:8080/PFE/result9.jpg

Figure III. 7:La page d'administrateur.

5.4 Connexion et inscription

Pour utiliser notre application, la première étape consiste en création d'un compte utilisateur. Une fois fait, l'utilisateur est demandé de saisir quelques informations personnelles comme le nom, prénom, filière, email, mot de passe ainsi que le pseudonyme. Toutes ces informations vont générer son profil d'utilisateur.

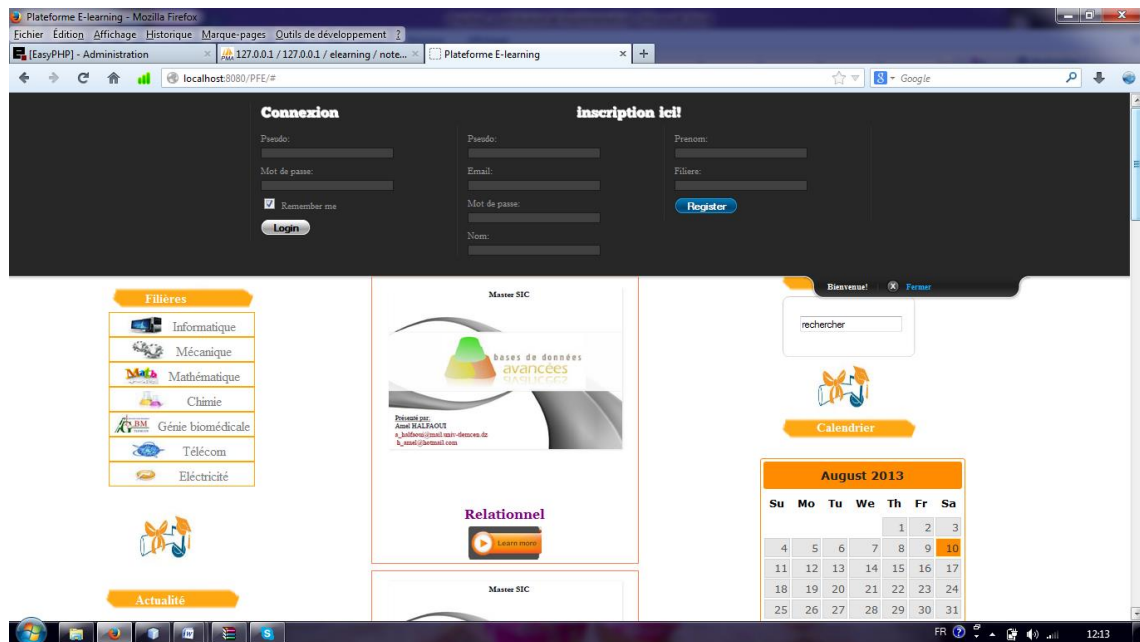


Figure III. 8: Connexion et inscription sur la plateforme.

5.5 Affichage et évaluation

Si l'étudiant décide de lire un cours il suffit de cliquer sur le bouton « Learn more », le système affiche la fenêtre ci-dessus et donne ainsi la possibilité d'évaluer le cours affiché, la note est comprise entre 1 et 10.

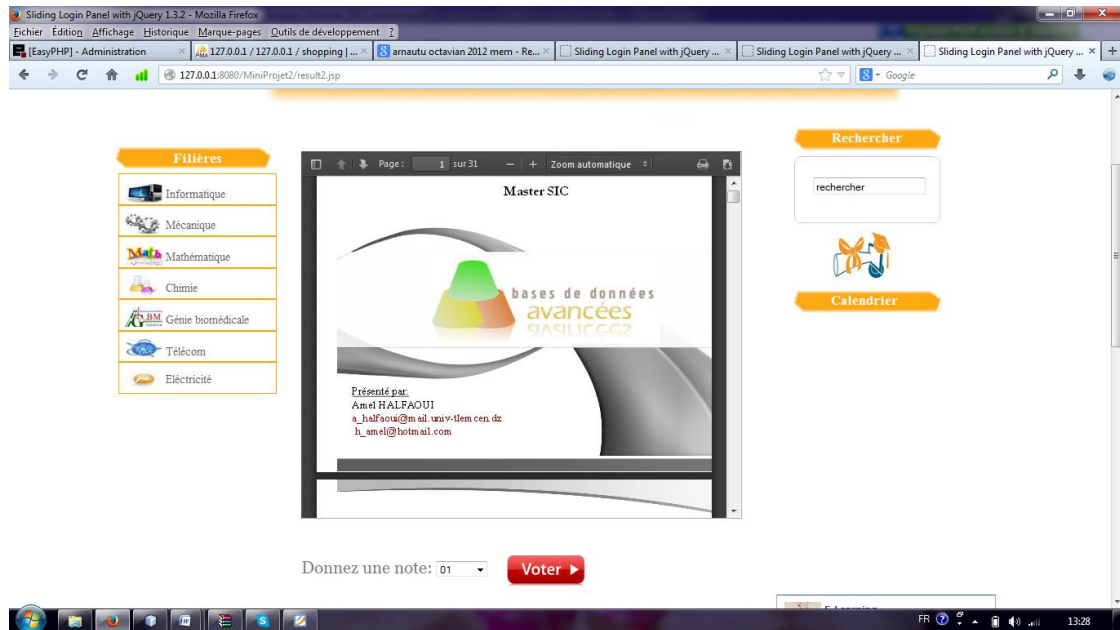


Figure III. 9:L'affichage et l'évaluation d'un cours.

5.6 Description de système de filtrage collaboratif

La figure III.10 présente les cours recommandés par le système à l'utilisateur actif après le lancement de l'algorithme de filtrage collaboratif :

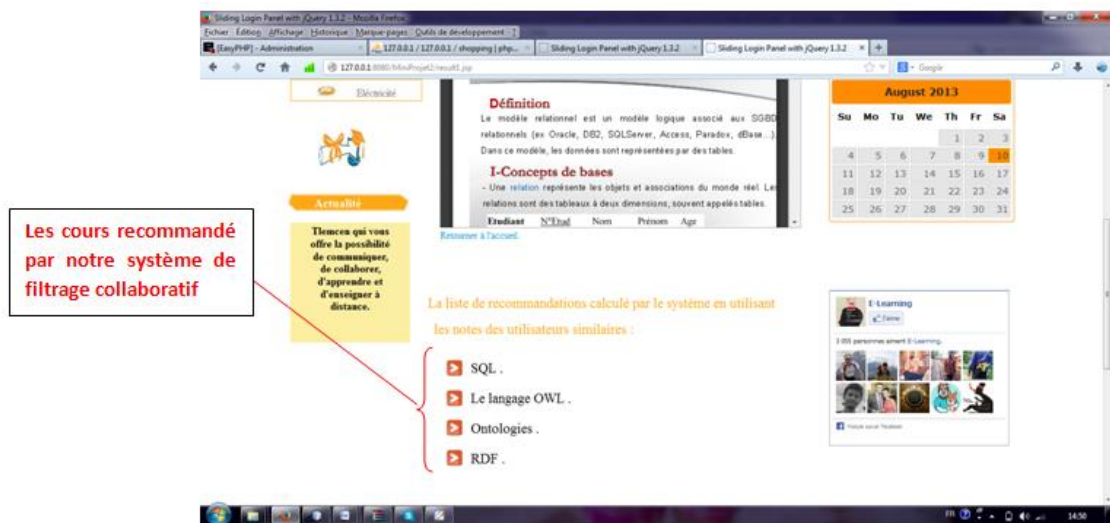


Figure III. 10:Recommandation des cours par le système.

Le filtrage collaboratif basé sur la mémoire utilise une matrice des votes contenant des préférences des utilisateurs pour prédire des ressources auxquels un nouvel utilisateur peut-être s'intéresse. Le FC basé sur la mémoire recherche les "k plus proches voisins" i.e. les k voisins les plus similaires à l'utilisateur actif, dans le but de générer des recommandations fiables.

Le principe général d'un système de filtrage collaboratif s'articule autour de deux fonctionnalités centrales :

- Le calcul de la proximité entre les utilisateurs.
- Le calcul de la prédiction de l'évaluation qu'un utilisateur fera d'un document.

5.6.1 Calcul de distances

	C1	C2	C3	...	Cj
U1	4		3	...	2
U2	5	4		...	
...
Ui	4	3	5	...	5

Tableau III. 1:Matrice des votes.

$\{C1, C2, \dots, Cj\}$:listes des cours.

$\{U1, U2, \dots, Ui\}$:listes des utilisateurs.

La base de données des votes d'utilisateur contient un ensemble des votes $v_{i,j}$ correspondant au vote de l'utilisateur i sur la ressource j . Soit I_i l'ensemble des ressources pour lesquelles l'utilisateur i a voté.

La moyenne des votes de l'utilisateur i est définie comme suit :

$$\bar{v}_i = \frac{1}{|I_i|} \sum_{j \in I_i} v_{i,j}$$

5.6.2 Calcul de prédiction

Les évaluations de l'utilisateur actif sont prédites à partir d'informations partielles concernant l'utilisateur actif, et un ensemble de poids calculés à partir de la base de données des évaluations des utilisateurs. La prédiction $P_{a,j}$ de l'évaluation d'une ressource j par l'utilisateur actif sera calculée par la formule :

$$P_{a,j} = \bar{v}_a + \kappa \sum_{i=1}^n \omega(a,i) (v_{i,j} - \bar{v}_i)$$

Avec :

- n le nombre des utilisateurs dont le poids n'est pas égal à zéro.
- $v_{i,j}$ l'évaluation de la ressource j par l'utilisateur i .
- \bar{v}_i la moyenne de l'ensemble des évaluations fournies par l'utilisateur i .
- $\omega(a, i)$ le coefficient de pondération liant l'utilisateur actif et i .
- K un coefficient de normalisation.

Le coefficient de pondération $\omega(a, i)$ représente la similarité existante entre l'utilisateur actif et les autres. Plus ils sont proches et plus le coefficient est grand. Ce coefficient peut être défini à partir de la distance, de la corrélation ou de la similarité entre chaque utilisateur i et l'utilisateur actif, la corrélation est calculée comme suit :

$$\omega(a, i) = \frac{\sum_j (v_{a,j} - \bar{v}_a)(v_{i,j} - \bar{v}_i)}{\sqrt{\sum_j (v_{a,j} - \bar{v}_a)^2 \sum_j (v_{i,j} - \bar{v}_i)^2}}$$

la valeur de $\omega(a,i)$ est compris entre -1 et 1. Si cette valeur est égale à -1, nous disons que les deux utilisateurs sont fortement corrélés opposés. Si la valeur est égale à 1, les deux utilisateurs sont fortement corrélés semblables. Les deux utilisateurs sont considérés être indépendants si la valeur est égale à 0.

5.6.3 Exemple de recommandation

En vue d'illustrer l'algorithme de Pearson, nous proposons un exemple très simple de la matrice de notes du tableau 2 qui représente trois apprenants pouvant noter quatre cours.

A partir des notes qu'ils ont attribuées aux cours, nous évaluons les similarités de notes entre ces apprenants.

✓ La matrice cours-apprenant

Supposons la matrice cours-apprenant suivante :

	Cours1	Cours2	Cours3	Cours4
apprenant1	7	-	8	7
apprenant2	6	5	-	2
apprenant3	-	2	2	5

Tableau III. 2:matrice cours-apprenant.

Chaque apprenant peut être vu comme étant un « vecteur » incomplet, on peut limiter notre analyse aux seules composantes qu'ils ont en commun.

Par exemple, pour comparer apprenant 2 et apprenant 3, on omet les valeurs manquantes ou les cours que l'apprenant2 et l'apprenant3 n'ont pas en commun. Nous avons alors les deux vecteurs suivants : $v=(5,2)$ et $w=(2,5)$.

La corrélation de Pearson est une mesure de la similarité entre deux vecteurs :

$$Pearson = \frac{v - \bar{v} \cdot w - \bar{w}}{\sqrt{(v - \bar{v}) \cdot (v - \bar{v}) * (w - \bar{w}) \cdot (w - \bar{w})}}$$

Où le produit scalaire (.) est calculé avec les cours que les deux utilisateurs ont en commun.

La valeur de Pearson est entre -1 et 1. Si cette valeur est égale à -1, nous disons que les deux utilisateurs sont fortement corrélés opposés. Si la valeur est égale à 1, les deux utilisateurs sont fortement corrélés semblables. Les deux utilisateurs sont considérés être indépendants si la valeur est égale à 0.

Dans le cas d'apprenant2 : $\bar{v} = (6+5+2)/3=4.33$, et d'apprenant 3 $\bar{w} = (5+2+2)/3=3$

Si on ne retient que les composantes correspondantes aux cours 2 et 4, les vecteurs sont donc

$(5-4.33, 2-4.33) = (0.67, -2.33)$ et $(2-3, 5-3) = (-1, 2)$ on a donc

$$Pearson = \frac{(0.67, -2.33) \cdot (-1, 2)}{\sqrt{(0.67, -2.33) \cdot (0.67, -2.33)} \cdot \sqrt{(-1, 2) \cdot (-1, 2)}} = \frac{-5.33}{5.43} \approx -1$$

En d'autres termes, l'apprenant2 et l'apprenant3 sont des « contraires » : l'un préfère le cours 4 et l'autre préfère le cours 2.

Pour calculer la corrélation entre l'apprenant1 et l'apprenant2, on observe d'abord que la moyenne d'apprenant 1 est 7.33. On a donc :

$$Pearson = \frac{(1.67, -2.33) \cdot (-0.33, -0.33)}{\sqrt{(1.67, -2.33) \cdot (1.67, -2.33)} \cdot \sqrt{(-0.33, -0.33) \cdot (-0.33, -0.33)}} = \frac{0.22}{1.33} \approx 0.16$$

En d'autres termes, l'apprenant2 et l'apprenant1 sont peu corrélés.

✓ Prédiction comme moyenne pondérée

Supposons maintenant que nous souhaitions prédire les notes que donnerait l'apprenant2 au cours 3. Nous utilisons alors la formule suivante :

$$\text{Apprenant2}(3) = \overline{\text{Apprenant2}} + \frac{\sum_{a=\text{apprenant1}, \text{apprenant3}} \text{Pearson}(\text{apprenant2}, a)(u_3 - \bar{a})}{\sum_{a=\text{apprenant1}, \text{apprenant3}} |\text{Pearson}(\text{apprenant2}, a)|}$$

Où $\overline{\text{apprenant2}}$ est la moyenne de toutes les notes accordées par apprenant2 (pour tous les cours) et même chose pour \bar{a} qui est la moyenne des notes accordées par l'apprenant a (sur tous les cours).

Ce type de formule est ce qu'on appelle une *moyenne pondérée*. Dans le cas qui nous concerne, on obtient la prédiction suivante :

$$\text{apprenant2} = 4.33 + \frac{-1 * (2 - 3) + 0.16 * 0.66}{1 + 0.16} \approx 5.3$$

Cet algorithme est l'algorithme de base du filtrage collaboratif apprenant-apprenant par corrélation de Pearson. Pour autant que tous les cours soient notés par au moins un apprenant ayant quelques cours en commun avec l'apprenant courant, on peut obtenir des prédictions efficaces pour tous les cours.

6 Conclusion

Au terme de ce chapitre, nous venons de conclure la dernière phase du projet, à savoir l'implémentation du système de filtrage collaboratif. Cette implémentation offre, via un ensemble de programme et de bases de données la possibilité de personnaliser un profile d'apprenant et de trouver les autres profiles similaires afin de recommander les cours pertinents à l'utilisateur actif.

Le filtrage collaboratif permet d'une part de résoudre un certain nombre de problèmes liés à la prise en compte des données multimédias ou des caractéristiques contextuelles des documents et d'autre part Les avis et conseils des collègues et des amis permettent souvent d'améliorer la qualité de la recherche personnelle.

Conclusion générale

Ce projet de fin d'étude représente le fruit de tant d'années d'études, il nous a permis d'avoir de nouvelles connaissances dans différents domaines.

La problématique abordée dans ce mémoire est centrée autour du problème de la recommandation des objets pédagogique sur un système e-Learning avec toutes les informations qui lui sont reliées.

Comme nous avons vu, l'objectif principal de notre système est le filtrage collaboratif qui permet de recommander à l'utilisateur actif des documents en se basant sur les évaluations des autres utilisateurs partageant les mêmes centres d'intérêt. Premièrement, nous avons calculé la similarité entre les utilisateurs. Ensuite, prédire des notes pour quelques documents et ne sélectionner que les documents avec un score élevé.

Tout le problème est donc de déterminer l'ensemble d'utilisateurs proches de l'utilisateur actif sur la base des jugements donnés en commun c'est-à-dire portés sur les mêmes documents.

En réalisant ce PFE, nous avons essayé de développer une mini plateforme e-Learning permet le recherche de l'information, le filtrage collaboratif des documents en utilisant la corrélation du Pearson l'un des algorithmes à base mémoire. Le système permet aussi de télécharger et lire le document, donné une note a ce document et recommander des documents à l'utilisateur actif en se basant sur les évaluations données par les utilisateurs similaire. Ainsi, il permet d'ajouter des nouveau cours à la base de données « cours ».

Perspectives

Néanmoins, tout projet de grande envergure, nécessite des efforts considérables et des améliorations continues, à cette occasion on note l'un des champs d'études dans le domaine qui restent à explorer :

L'usage des concepts du web sémantique dans le filtrage d'information collaboratif pour la modélisation d'une plateforme E-Learning, cet axe peut diviser en deux objectifs :

Le premier objectif consiste à implémenter deux ontologies :

Conclusion générale

- La première est pour générer les profils utilisateurs en utilisant le vocabulaire FOAF (Friend Of A Friend) qui permet de décrire les personnes, leurs comptes, ce qu'elles font, les liens entre elles, l'appartenance à des groupes.
- La deuxième est pour modéliser les objets pédagogiques en utilisant le standard IMS-LD qui permet d'assurer l'accessibilité des objets pédagogiques, en permettant de les décrire et les retrouver, de les adapter à des profils d'utilisateurs.

Le deuxième objectif : consiste à réaliser un système de recommandation collaborative des objets pédagogiques qui utilise ces deux ontologies pour donner un aspect sémantique au système.

Les références

- [1] Nguyen An-Te, COCoFil2: Un nouveau système de filtrage collaboratif basé sur le modèle des espaces de communautés, 2006, Thèse de doctorat, Université Joseph-Fourier-Grenoble I.
- [2] Belkin Nicholas J ET Croft W. Bruce, Information filtering and information retrieval: two sides of the same coin? , Communications of the ACM, 1992, vol. 35, no 12, p. 29-38.
- [3] Oufaida Houda ET Nouali Omar, Le filtrage collaboratif et le web 2.0. Etat de l'art, 2008, Vol. 11 p.13-35.
- [4] Belloui M. Amokrane, L'usage des concepts du web sémantique dans le filtrage d'information collaborative, 2008, Thèse de magistère, Institut National d'Informatique Alger.
- [5] Peis Eduardo, Del Castillo JM Morales, et Delgado-López J. A., Semantic recommender systems analysis of the state of the topic. Hipertext.net, 2008, vol. 6, p. 1-5.
- [6] Woerndl Wolfgang ET Groh Georg, Utilizing physical and social context to improve recommender systems, In : Proceedings of the 2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Workshops, IEEE Computer Society, 2007, p. 123-128.
- [7] <http://www.grappa.univ-lille3.fr/~candillier/profils/dea002.html>, Consulté le 03.05.2013.
- [8] Berrut Catherine, « Filtrage collaborative », chapitre 8, p 255-283, E. GAUSSIER, M.H. STEFANINI, « Assistance intelligente à la recherche d'information »,Hermès-Lavoisier ,2003
- [9] http://fr.wikipedia.org/wiki/Corr%C3%A9lation_de_Spearman, Consulté le 03.05.2013.
- [10] L. Candillier, “Apprentissage automatique de profils de lecteurs”, 2001
- [11] Heckerman David, A tutorial on learning with Bayesian networks.1999

Références

- [12] Castagnos Sylvain, Modélisation de comportements et apprentissage stochastique non supervisé de stratégies d'interactions sociales au sein de systèmes temps réel de recherche et d'accès à l'information, 2008, Thèse de doctorat, Université Nancy II.
- [13] Pennock David M., Horvitz Eric, Lawrence Steve, et al. Collaborative filtering by personality diagnosis: A hybrid memory-and model-based approach, In: Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., 2000. p. 473-480.
- [14] Zaïer Zied, Modèle multi-agents pour le filtrage collaboratif de l'information, 2010.
- [15] Iles Nawel, Chikh Azeddine, Chouiti Sidi Mohammed, Architecture d'un entrepôt pédagogique et sa modélisation sémantique, International Conference on Web and Information Technologies ,2008.
- [16] Nguyen Thanh Binh, Gestion de profil apprenant distribue dans un environnement ubiquitaire, 2009, Mémoire de stage de fin d'études, Institut de la Francophonie pour l'Informatique.
- [17] <http://www.imc.org/pdi/>, Consulté le 08.05.2013.
- [18] [<http://www.w3.org/TR/vcard-rdf>], Consulté le 08.05.2013.
- [19] Dumbill Edd, XML Watch: Finding friends with XML and RDF: The Friend-of-a-Friend vocabulary can make it easier to manage online communities, 2002.
- [20] <http://fr.wikipedia.org/wiki/Amazon.com>, Consulté le 11.08.2013.
- [21] http://fr.wikipedia.org/wiki/Filtrage_collaboratif, Consulté le 11.08.2013.
- [22] Denos Nathalie, Berrut Catherine, Gallardo-López Lizbeth, et al. COCoFil: une plateforme de filtrage collaboratif orientée vers la communauté. In: CORIA. 2004. p. 9-26.
- [23] Goundo Camara Fatoumata ET Guillaume Argoud, COCoFil3: Community-Oriented Collaborative Filtering, 2007.
- [24] Castagnos Sylvain, Boyer Anne, Charpillet François, et al. Vers un filtrage collaboratif distribué: le modèle rsb. In: Modèles Formels de l'Interaction (MFI'05). 2005. p. 260-268.

Références

- [25] Berkovsky Shlomo, Eytani Yaniv, Kuflik Tsvi et al., Hierarchical neighborhood topology for privacy enhanced collaborative filtering. In: Proceedings of PEP06, CHI 2006 Workshop on Privacy-Enhanced Personalization, Montreal, Canada, 2006. p. 6-13.
- [26] Zaïer Zied, présentation du projet de recherche : Modèle multi-agent pour le filtrage collaboratif de l'information, 2003.
- [27] Zaïer Zied, Modèle multi-agents pour le filtrage collaboratif de l'information, 2010.
- [28] Boussebough Imane, les systèmes multi-agents dynamiquement adaptables ,2011.
- [29] Mazyad Hanaa, Une Approche Multi-agents à Architecture P2P pour l'Apprentissage Collaboratif, 2013.
- [30] Karen HL TSO-SUTTER, Leandro Balby MARINHO ET Lars SCHMIDT-THIEME, Tag-aware recommender systems by fusion of collaborative filtering algorithms. In: Proceedings of the 2008 ACM symposium on applied computing, 2008. p. 1995-1999.
- [31] Szomszor Martin, Cattuto Ciro, Alani Harith, et al., Folksonomies the semantic web and movie recommendation, 2007.
- [32] Bouzidi Driss, Ajhoun Rachida, Benkiran Amine et al. Les annotations pour un apprentissage collaboratif. In : Premières journées communication et apprentissages instrumentés en réseau. 2006. p. 372-382.
- [33] <http://pages.univ-nc.nc/~touraivane/Java/node4.html>, Consulté le 13.08.2013.
- [34] <http://fr.wikipedia.org/wiki/NetBeans>, Consulté le 13.08.2013.
- [35] <http://www.journaldunet.com/encyclopedie/definition/549/34/20/struts.shtml>, Consulté le 13.08.2013.
- [36] http://tahe.developpez.com/java/struts2/?page=page_1, Consulté le 14.08.2013.
- [37] http://www.microapp.com/definition_easyphp_132.html, Consulté le 14.08.2013.
- [38] <http://www.learnperfect.fr/elearning/>, Consulté le 12.08.2013.

ANNEXE A

La Figure A.1 représente le code pour calculer la corrélation de Pearson

```
public class Correlation {  
  
    public static double[] x =new double[4];  
    public static double[] y =new double [4];  
    public static double[] z =new double [4];  
    public static ArrayList <Double>resl= new ArrayList();  
    public static Connection con;  
    public static int cont;  
    public static double r;  
    public static double mean_xl=0,mean_y1=0;  
  
    String login;  
  
    public void setLogin(String login) {  
        this.login = login;  
    }  
  
    public String getLogin() {  
        return login;  
    }  
  
    public static double pearson (double[] x, double[] y) throws SQLException{  
        int nb=0, nbx=0,nby=0;  
  
        assert x.length == y.length;  
        double mean_x = 0;  
        double mean_y = 0;  
        int n_x = x.length;  
        int n_y = y.length;
```

Figure A. 1:Le code de corrélation de pearson

La Figure A.2 représente le code pour calculer la prédiction :

```
for (int i = 0; i < z.length; i++){  
    if(z[i]>0){moy_y += z[i];  
        nby=nby+1;}  
}  
  
moy_y /= nby;  
  
for (int i = 1; i <=x.length; i++){  
    if((x[i]==0)&&(z[i]>0)){  
        double pred=mean_xi+((maxx*(z[i]-moy_y))/Math.abs(maxx));  
        int k=i+1;  
        System.out.println("pred de cours "+k+" = "+pred);  
    }  
}  
}  
} catch(Exception e) {  
}
```

Figure A. 2:Le code de prédiction

Résumé

Les systèmes de recommandation deviennent, indubitablement, de plus en plus populaires. Cette popularité est le fruit de l'assistance fournie par ces outils aux utilisateurs dans leurs recherches, les rendant à la fois plus personnalisées et plus optimales. Le filtrage collaboratif est l'une des approches les plus répandues des systèmes de recommandation.

Dans ce mémoire nous nous intéressons à l'utilisation du filtrage collaboratif dans une plateforme e-Learning. Notre objectif est de recommander des objets pédagogiques à l'utilisateur actif, le principe est d'utiliser l'ensemble de la base de données des évaluations des utilisateurs pour identifier les utilisateurs voisins (les plus proches), et faire prédire les évaluations de l'utilisateur actif afin de lui recommander les ressources pertinentes. Plusieurs mesures ont été exploitées dans le cadre du filtrage collaboratif parmi ces mesures nous avons utilisé la corrélation du Pearson.

Mots-Clefs : Les systèmes de recommandation, filtrage collaboratif, e-Learning, utilisateur actif, les utilisateurs voisins, corrélation du Pearson.

Abstract

Recommendation systems become undoubtedly more popular. This popularity is a result of the assistance provided by these tools to users in their searches, making them more personalized and optimal. Collaborative filtering is one of the most popular recommender systems approaches.

In this paper we are interested by the use of collaborative filtering in an e-Learning platform. Our objective is to recommend learning objects to the active user, the principle is to use the database of user's ratings to identify neighboring users (the closest), and predict the evaluations of the active user in order to recommend him the relevant resources. A lot of measures have been used in the context of collaborative filtering among these measures we used the Pearson correlation.

Keywords: Recommendation systems, collaborative filtering, e-Learning, active user, neighboring users, Pearson correlation.

ملخص

نظم التوصية أصبحت بلا شك أكثر شعبية، هذه الشعبية هي نتيجة المساعدات التي تقدمها هذه الأدوات للمستخدمين في عمليات البحث الخاصة بهم، مما يجعلها على حد سواء أكثر تخصيصاً والأمثلة التصفية التعاونية هي واحدة من أساليب نظم التوصية الأكثر شيوعاً.

في هذه المذكرة نحن مهتمون باستخدام التصفية التعاونية في نظام التعليم الإلكتروني. وهدفنا هو اقتراح دروس للمستخدم الناشط، المبدأ هو استخدام قاعدة البيانات الخاصة بتقييمات المستخدمين من أجل تعريف المستخدمين المجاورين (الأقرب)، و

تخمين تقييمات المستخدم الناشط من أجل اقتراح الدروس الموافقة له. مجموعة من القوانين استخدمت في سياق التصفية التعاونية من بين هذه القوانين استخدمنا ارتباط بيرسون.

الكلمات المفتاحية: نظم التوصية، التصفية التعاونية، التعليم الإلكتروني، المستخدم الناشط، المستخدمين المجاورين، ارتباط بيرسون