Contextual Information Retrieval within Recommender System: Case Study E-learning System

M. Brik, M. Touahria

Abstract
Thisis paper focuses on monitoring and analyzing user activities on CF-based recommender system in order to guess suitable and unsuitable items' context information using rating matrix which making more efficient adaptation task. An ontology-based user profile and rules-based context modeling for reasoning about context information is proposed in this research work, in addition to an investigation to apply Semantic Web technologies in user modeling and context reasoning. This proposal is applied in education field in which we have designed an authoring tool for learning objects within ubiquitous environment. This system aims to improve the learning object production task (creation, review, edition…) on behalf of technologies offered by collaborative filtering systems as well as user behaviors monitoring to improve the recommendation process.
Keywords:
Collaborative filtering, user profile, context aware, rule-based ontology, user behaviors.

1. Introduction

Recommender systems (RS) have obtained significant importance in the last decade which provides a relevant data source (rating data). This paradigm has been used in many domains, such as E-commerce, where the recommender systems are used to provide different products to customers with different needs. In tourism area these systems are used to retrieve personalized and appealing location and objects for the potential users of touristic products.

Data generated by recommender engine are used to construct a decision support model. The RS will offer an amount of information easy to manage, adapted to the user needs and preferences. An important topic has been extensively used in recommendation system is called Collaborative filtering (CF). This last one used a rating matrix which is the basis of recommendation generation in CF-based recommendation system that contains both rated and predicted data value. A rating score is given directly by user of interest while a predicted value is offered by the system using data mining methods. Systems that are able to detect the context in which users operate the items were exposed to enhance the recommendation method. This paradigm exploits different methods to recognize the effect of contextual information on prediction of the ratings value. These systems are called CARS (Context Aware Recommended System) that integrate the context aspect into recommendation mechanism to generate more personalized objects and services. Contrariwise, recommender system which do not take the context aspect into account maybe lost in predictive task.

Analysis of users' interactions with the items provides important information about users' behavior, a behavior which is defined as a concept that models the characteristics of a user interacting with a system [22] provides important information on the consumption of context resource. A user behavior monitoring and analysis is an important way that aids to generate implicit data and can be fully used to make the system adapted to the user. It has been used by lot of systems that support recommendation, this work intents to analyze the user behavior in ubiquitous environment in order to deduce relevant information relatively to the resource context. Our system has been designed for this purpose and allows to retrieve relevant and irrelevant context information by analyzing the interaction of the user with the resource, because the user interaction reflects user's behaviors and interests. In another way our system answers the question from all contextual data that can be acquired, what are suitable and what are not suitable for a specific resource and how it can be use afterwards? As a response for this question we have applied this work on a dedicated recommender system for e-learning for which we propose an authoring system within users' community. All our users were considered as authors with different levels (beginner, expert, professor, lecturer ...) and we put an assumption that the users utilize different devices equipped with different configuration (smart phone, PC). The collaborative filtering techniques is the platform of our work and we analyze the user behavior inside collaborative filtering system taking into account the time spent on learning object and a collaborative filtering result set.

Semantic-based technology offers the way to modeling user and its interactions. The ontological model gives many advantages [12] which is enabled the representation of semantic information and permit reasoning via semantic-based rules which can enrich the representation by inferring unknown facts. On other hand, enriching user profile data with semantic context information is useful to infer knowledge about what is the requirement in the adaptation process. The context of user interaction presented in this work is composed of three portions as indicated in [29]. These portions are environment, user and platform. User is described by its competencies and demo-

Manuscript received January 5, 2026
Manuscript revised January 20, 2026
https://doi.org/10.22937/IJCSNS.2026.26.1.7

graphic information. Platform is the set of hardware (devices) that intervene in the interaction. Environment refers to the set of pieces that user interacted with it (learning objects for our application). The second benefit of our designed system is its ability to present an authoring system for novice author (like beginner lecturer, author …) who needs to know the point of view of her/his users community about her/his learning object being created by addressing the query to the subset of author's community (considered as expert authors, professors,…) in order to know their opinions (rating data value) about the learning object. This proposal aims to help the author to improve his/her learning object taking into account the opinions of all collaborators. This application focuses on the recommended performance in memory-based collaborative filtering algorithms. The core of collaborative filtering is to calculate similarities among authors and learning objects documents.

2 RELATED WORK AND MOTIVATION

The most existing approaches that are used in acquisition context were based on explicit, implicit and/or inferred contextual data [30] used physicals sensors (GPS, RFID …). In [4], [24] inferred automatically the device characteristics in order to calculate the suitability or likeability of applicant device. Other works has been based on manual resource description which can adjust or describe what are then context information is suitable for. The work in [3] presents a device capabilities detection (screen size, resolution) for adaptable user interface, this approach is based on fuzzy-reasoning mechanism to infer new user and device capabilities. In previous approaches it is noted that the context suitability decision is restrained to the resource holder whose resource context value required is difficult to be precisely defined, which leads sometimes to mistaken adaptation process. Our approach is different as it solved the problem on the client side i.e. the user interactions with resource helps us to infer the appropriate context information.

A user interaction has been studied in many works, for instance , in [25] a user profile data has been automatically extracted using users' community topics detection to infer relevant resource context information, [2] proposed a method that computes customized recommendation by combining past behavior of user and user community behavior. Many other works has been proposed ontologies in order to describe the context of human activities. We found in [23] the most relevant works organized according to context parameters (location, time, user preferences ...). A user's preferences ontology that describes device capabilities is used in [33]. The representation model can guide the adaptation of the content taking into account the device characteristics. The study in [5] presents a survey for semantic-based context reasoning approach, this work also listed many various context aware systems and tools that incorporate ontologies. The authors in [7] have described the SOUPA ontology (standard ontology for ubiquitous and pervasive ontology) written in OWL (ontology web language) for the purpose to modeling context in pervasive environment. Other example CANON [31] an ontology for modeling context in pervasive computing environment that presents a context model and logic-based context reasoning schemes, in this work a context reasoning was focused on location (bedroom, bathroom, kitchen, …) to derive user's situation in smart phone scenarios. Other work has extended the CANON ontology by integrating a temporal ontology and rules-based context aware smart home [32]. Five rules are presented in [8] for multimedia conferencing process according to the user notification services (Email, SMS, voice) and conferencing time efficiency, this strategy was implemented using rules language defined by JENA framework.

Other many works are tailored a rule-based model for modeling and reasoning their context, we refer the reader to [23] for more examples. Above all we believe that the use of semantic model provides a very powerful way to describe items and their relationships of users' profile which improves the effectiveness of recommendation task, the main contribution in this paper: We defined a model for user profile that includes environment such as devices, items characteristics (learning objects in our use case) and inferences rules that modeling the user behaviors in order to retrieve relevant and irrelevant context information.

We show how to utilize the retrieved information and we apply this proposal in education field in order to improve the recommendation task. We tailored a collaborative filtering system to suit our needs and we have added two new metadata elements to the L.O.M (learning object metadata) schema which can be automatically filled in order to store and manage the retrieved information. The rest of this paper is organized as follows. Section 2 briefly describes the background regarding recommender systems. Section3 describes the user profile and inference rules. The detailed description of our system can be found on section 4. The evaluation and experimentation results are presented on section 5. Finally, section 6 is devoted to summarize the conclusions and future work.

3 KNOWLEDGE BASE FOR RATING DATA

The most relevant things in collaborative filtering based recommender systems is rating matrix which rows represent users and the columns represent items, this matrix can be used to infer latent information related to the user preference. In fact, when the user rated a specific item with high score implies that the end user has consumed the item with confortable context. The knowledge base used in this study is composed of three layers: scores layer, user attention layer and items layer. The score layer represents the possible score given within recommender system (high, low, none), item layer represents item characteristics and user attention layer symbolizes potential user cognition state regarding an item. The Figure 1 illustrates the possible rating data in which we have supposed a threshold that separate data into two categories (high and low). The use of time counter aids to know the time spent on item which helps us to know the user attention

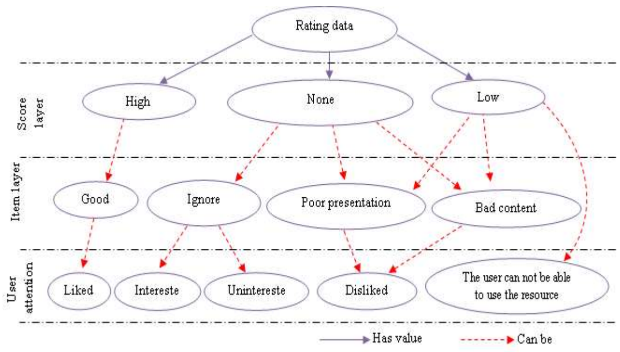


Figure 1. The possible rating data into two catego

(interested or uninterested). consequently, our knowledge base represents the facts about rating score within collaborative filtering system and possible causes of generation (which is not exhaustive). As shown in Figure, when user liked an item that mean that the user is comfortable with it, in that event, we have considred that the item context is suitable for the user. Contrary, when the user disliked an item there are several reasons caused by (as shown in above Figure). Our approach is based on two assumptions, first: a high score given by a user implies the user context is appropriate and the second: in some cases, the abstaining from rating an item is caused by the incompatibility of device resources with the item content.

3.1 User profile

A user profile is a set of information that characterizes a specific user which such recommender system can use it to perform the adaptation task. Generaly a user Profile is represented as a set of weighted keywords, semantic networks, weighted concepts, or association rules. The most common description for user profiles is set of keywords which can be automatically extracted from documents and/or provided by the user itself. The construction of user profile is based on information sources, using a diversity of construction methods such as information retrieval or machine learning [1]. the user profile in our use case contains a set of weighted keywords for characterizing user competencies and items (keywords-based items classification), some detailed information about user's community like demographic information, interests, and competencies for identifying a user and the hardware device characteristics, user interaction with items and history are also a part of the user profile. For the first time, a user must complete a questionnaire about the personal information and competencies, afterward any activity imply a recalculation of user competencies using some predefined rules, and finally user profile will be restructured automatically after any change in user history.

3.2 Rules based context reasoning

The main contribution in this work is detecting suitable and unsuitable context information using rating data provided by recommender engine. The user behaviors recognition with consideration of user session duration and data rates offer an important way to predict the suitable and/or unsuitable context information that depicted by a set of information about hardware resource, which allows us to make recommendations for target user taking into account all retrieved information.
The strategy that we have applied in order to accomplish our task is based on two major criteria: one is time spent on item, and the second global rate of item provided by recommender engine.

Rule-based reasoning is a powerful method allows us to derive relevant contextual information and relatively easy to implement using data provided by sensors, the information acquired from context sensors cannot be directly used for adapting arbitrary item. Therefore, useful contextual information can be obtained from context data according to a set of rules defined for each item. Through

Rule1 (table below), our system is capable to determine the ability of user competences that participate in rating process. k represents user competencies as list of keywords and k' represent the item classification as list of keywords, the built-in swrlb:Listintersection is used in order to know the common keywords between user and item, it satisfied when the intersection between list keywords (k) and list keywords(k') is not empty.

Rule2 aims to determine the user attention (interested or not). The user is interested by an item when he/she has the ability to rate item and spends enough time on item. By *Rule3* our system is able to detect the suitability of user context, this rule is based on fact that user who scores the item with high score signify the user has appropriate context. *Rule4* aims also to infer the user attention about an item (ignored), this rule is based on a time counter, whether the user did not spend enough time t on the item which it require time t' we infer that the user is ignored the item. Contrary, *Rule5* provides us the set of uninterested users. *Rule6* and *Rule7* aim to elicit the user's competencies as keywords list. Finally *Rule8* aims to retrieve the unsuitable context value which is based on second assumption discussed above. This rule consider that if a user do not rate the content and spends a sufficient time on the item and if her/his predicted score equal "high" and the final score for item equal "high" then we decide that her/his context is not suitable.

TABLE 1 RULES-BASED CONTEXT REASONING

ID	RULE
R1	$hasKeywords(?u,?k)$, $hasKeywords(?it,?k')$, $swrlb{:}\,listIntersection(?k',?k) \rightarrow$ $hasAbilityToRate(?u,?it)$
R2	$hasAbilityToRate(?u,?it), hasContext(?u,?c), spend(?u,t),$ $require(?it,t'), swrlb{:}\,greaterthan(t,t') \rightarrow$ $interest(?u,?it)$
R3	$interest(?u,?it), rate(?u,\text{"HIGH"}), hasContext(?u,?c) \rightarrow$ $hasSuitableContext(?u,?c)$
R4	$hasAbilityToRate(?u,?it), spend(?u,?t), swrlb{:}\,lessthan(t,t'),$ $require(?it,t'), rate(?u,\text{"none"}) \rightarrow$ $ignoreWithoutInterest(?u,?it)$
R5	$hasAbilityToRate(?u,?it), spend(?u,?t), swrlb{:}\,greaterhan(t,t'),$ $require(?it,t'), rate(?u,\text{"none"}) \rightarrow$ $ignoreWithInterest(?u,?it)$
R6	$Create(?u,?it), hasKeywords(?it,?k) \rightarrow$ $hasCompetence(?u,?k)$
R7	$rate(?u,\text{"high"}), hasKeywords(?it,?k') \rightarrow$ $hasCompetence(?u,?k')$
R8	$interest(?u,?it), rate(?u,\text{"none"}), predictedScore(?it,\text{"high"}),$ $globalscore(?it,\text{"high"}), hasContext(?u,?c) \rightarrow$ $hasUnsuitableContext(?u,?c)$

4. CASE STUDY IN THE EDUCATION FIELD

4.1 System overview

By the following, we describe our system which is a tool for authoring purpose in education filed, it allows users to create new learning object and/or evaluate multimedia learning objects created by other users. The proposed system has two benefits. First, it is intended to help users to create new learning objects by providing a collaborative environment, in which interested users can participate in content assessment. Users who participate in this mission can 42 some problems caused by different context configuration (resources hardware), so to solve, we are obliged to take charge of the context configuration in future distribution of this object, this represents the second advantage of this system, which we have tried to determine the appropriate and inappropriate context data according to the score provided directly by the users or pridected by the system as well as their behaviors. r systems consists of five components: (a) Metadata extractor, (b) Document similarity calculator, (c) Users potential filtering, (d) Rating and predicting missing data manager, whose functions are elaborated below

a) Metadata extractor: this module is responsible for fulfilling the metadata elements, it shows an interface to fill all needed information which can be automatic like our proposed elements (discussed later) and all information's (date, time, size…) that can be gathered automatically or semi automatic like keywords list generated automatically using formula (detailed below), or manually like document name, etc.

b) LOs similarity module: this module aims to find similar LOs from system's database applying a cosine similarity approach using tf-idf weighting approach, although all documents has been presented as vector weighted in order to apply this formula

c) Users potential filtering: this module aims to retrieve a set of similar user based on K-Nearest Neighbor algorithm using the Pearson correlation coefficient and the keywords list generate by above module and attempt to send the LO to this set of users in order to invite them to give their rate about the LO being created .

.d) Rating and predicting missing data: this module is respon sible for collecting the rate from similar users and predicts all missing data in order to calculi the average between them, it use the LOs similarity module and user similarity module to perform the predicting task.

Finally all data (learning object and its metadata) are stored in a database for further access by students, lecturer and authors.

4.2 Ontology based user profile

In this study, we adopt the model represented by ontology, which allows us to represent the model using standard computer languages like OWL and modeling the elements of a structured context. The ontology is a formal specification of concepts and terms and relations between them [1]; it allows us to represent formally the dependencies between the different components of the context. In the present use case, the context kind is represented by bandwidth and support multimedia hardware (image quality, screen resolution), also our ontology includes user characteristics and interactions, items characteristics and recommender system data aspect.

Our goal was twofold. We firstly tried to define the conceptual vocabulary mobilized for the representation of knowledge in communities of the authors of educational resources. On the other hand, we also wanted to reuse the ontology of the domain of rating educational resources proposed in the literature by integrating them.

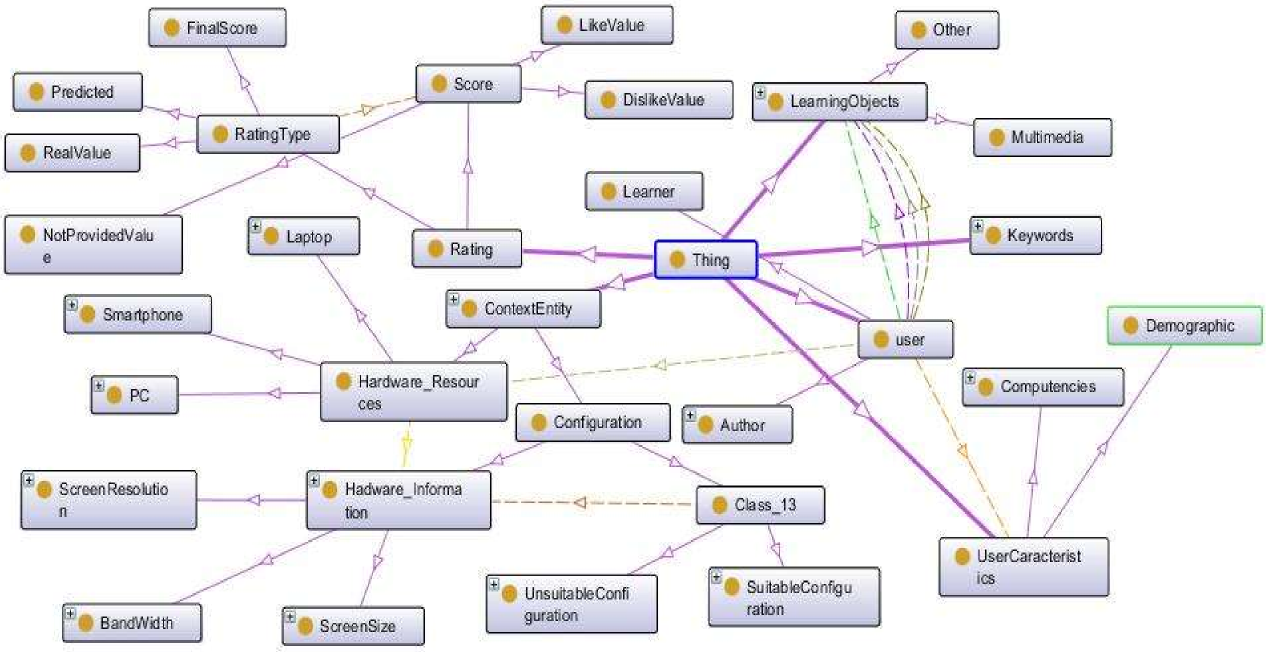


Figure 1 user profile ontology

Describing learning objects

Metadata standards

In many research domains, the most common way to describe an object is to use metadata; these descriptors are significant in the education field for access, retrieve and reuse the learning object. The present work uses a set of metadata attributes (metadata schema) in order to describe the user context and its environment also describing and indexing the learning objects. A learning object is a sort of digital element that permits content reuse, independence and flexibility in order to give a high quality of control to users [32]. However to get better learning object description, the use of metadata is necessary to accomplish this task. The common definition of Metadata is data about data; therefore to ensure interoperability with other systems we must used a standard. By the following we give details of standards used in educational field.

The Dublin Core (DC), invented by Dublin Core Metadata Initiative (DCMI) Is a simple metadata schema which is used in many work [11] this schema is presented as a set of 15 feature (Title, Identifier, Language, and other), the main key to use this schema is that compatible for all domains, furthermore many other additional attributes are invented called qualifiers that refine the 15 base elements to increase the efficiency of learning object indexing for more details, we refer the reader to [11].

(IEEE) Institute of Electrical and Electronics Engineers invented a dedicated standard for education context that allows the effective learning object description, this metadata schema is used in many LOR (learning object repository), called IEEE 1484.12.1-2002 Learning Object Metadata Standard (LOM) [15]. This schema provides categories and each category contains some elements and thus, in whole, LOM offers 76 data elements.

Metadata construction phase

The context information kind studied in this use case seems useful for an appropriate distribution of learning objects. In order to retrieve the suitable context information, we need to collect and store the context data used in rating phase for each participant (screen size, screen resolution and internet bandwidth), so to accomplish this task, we propose to add an extension to the LOM standard, this extension aims to preserve interoperability with other educational systems and also facilitate the adaptation treatment, to achieve this, we refer to [6] when he proposed an extension of LOM to MLM Mobile learning metadata that consist of 3 top level categories:1) Learning object which consist of information describing the learning resource, 2) Learner which consist of information describing the learner 3) setting which consists of information describing the context state of the learning environment so, in our work we use the standard LOM to describe the learning object and the extension proposed is Suitable_Context and Unsuitable_Context at technical category (branch 4.4.1.5 and 4.4.1.6) relatively to suitable configuration recommended for using this learning object which calculate automatically using rule based approach, and minimal configuration (unsuitable) required i.e. the context information of end user must be greater than for using rightly the learning object.

TABLE 2 PROPOSED METADATA ELEMENTS

CATEGORY	ELEMENTS LOM	SUB ELEMENT
4- TECHNICAL	4.4.1.5 SUITABLE_CONTEXT	4.4.1.5.1 NAME
		4.4.1.5 .2 VALUE
	4.4.1.6 UNSUITABLE_CONTEXT	4.4.1.6.1 NAME
		4.4.1.6.2 VALUE

Generation of metadata elements

In order to describe the learning content about the subject covered we have designed and implemented an extracting keywords algorithm. The most used formula in this context is the weighting term frequency – inverse document frequency (tf-idf). To use (tf-idf) the document must passes thought many phases, like Tokenization (sentences are splitting into words) and Remove Stop-word (i.e. words that haven't any meaning for the subject) and finally Stemming (using a specific morphologic analysis related to current language, each word is abridged to its morphologic root)

$$w_{ij} = tfidf(t_i, d_j) = tf(t_i, d_j) * log \frac{|D|}{tf(t_i, D)}$$

Where tf(ti,dj) represents how many time the term ti appear in document dj (term frequency (tf)); |D| is the number of documents in the corpus; tf(ti ,D) refers to the number of documents in the corpus that term ti appear in.

As a result of this phase we obtain an ordered vector representation of the document dj as a vector of (term| weight).

$$d_j = \{(t_1|w_1), (t_2|w_2), (t_3|w_3), ..\}$$

Where $w_1 > w_2 > w_3 > \cdots$

The result is sorted according wi in order to give the N first words (Top-N) that are candidate as keywords for the document. Our system provides the possibility to authors to change, edit or extend the keywords list given by system in order to overcome some limitations recognized by TF-IDF approach's [21],[14]. The following example shows the metadata encoded in XML [19].

```
<lom:general>
<lom:title>
<lom:string language="en">
        Title of the Learning Object
</lom:string>
</lom:title>
<lom:language>en</lom:language>
<lom:keyword   weight =”0.34”>
<lom:string language="en">Keyw_1</lom:string>
</lom:keyword>
<lom:keyword   weight =”0.28”>
<lom:string language="en">Keyw_2</lom:string>
</lom:keyword>
</lom:general>
```

Learning object rating phase

After the construction of metadata, our system accesses to the user database to find a set of similar users in order to collect their score on learning object being created, the purpose of this idea is to benefit of authors' experiences in order to get a final score of learning content. To achieve this, we refer to the recommendation systems technology which provides relevant techniques used by this work. In field of technology-enhanced learning (TEL) there are many works focused on recommendation system to retrieve suitable and pertinent learning object to the end-user (students), in [28] applying collaborative filtering directly to matrix user-rating in context of recommending music, a system have been proposed for the recommendation of learning resources, it integrate a collaborative filtering module that operates with ratings offered by users and equipped with inference rule engine, another study is the LORM tool (Learning Object Recommendation Model) [27] it use a hybrid method that recommends a preference-based and correlation-based learning objects for learners, this tool agreed an ontological model to performing semantic discovery. as summarize, the most rating-based systems for learning object manipulation was concentrated solely on the standpoint of learner i.e. the feedbacks returned by learner are used to improve the learning object, however this present some limitations because the learner makes comments on what he/she sees in content but in the case of a shortage or lack of reference or something important learner could not be able detect this lack in the majority of cases. Many other works are based on recommender system technique to deliver the suitable learning content, we find that the most of this system are focused on learner activity which we are discussed the disadvantages in the above section. We find in [17] a review of the most recommender system focusing on teachers (as expert community)

Learning objects similarity module:

In literature the cosine similarity [9] is frequently used when trying to determine similarity between two documents which the document is represented as vector and the cosine similarity calculate the inner product space that measures the cosine of the angle between them and the range of resulting similarity is between -1 and +1. Giving two documents A and B the cosine similarity between A and B is:

$$\text{Similarity} = \frac{A.B}{\|A\|.\|B\|} = \frac{\sum_{i=1}^{n} A_i \ . \ B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2}\sqrt{\sum_{i=1}^{n}(B_i)^2}}$$

Our approach calculates the similarity between given document and all LOR documents by using Cosine Similarity which is used in order to recommending a subset of LOR documents that consider as pertinent. User potential filtering. After pre-processing and weighting learning object, the next step is to collect all rating data about learning object being created from users whose jugged efficient to rate this object in order to calculate the average rating score. Our choice for giving a teacher's cluster is the K-Nearest Neighbor algorithm.

Known as user-user collaborative filtering, K-Nearest Neighbor is a supervised learning algorithm, which is the most common method used for prediction, estimate, and classification [10],[20]. We need for this algorithm in order to give predictions for learning objects for each user that has not rated the object. The process of this phase is as follow:

1. Calculate the similarities between active user (T1) and all users (Tj)
2. Select N top users given by step 1. (N represents the max number of selected user)
3. Calculate the prediction for the learning object.

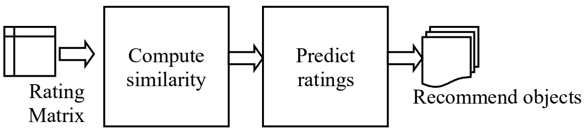


Figure 3: Recommendation process

One of success method of similarity measures used in collaborative filtering field is the Pearson Correlation Coefficient (PCC) [10] which measures the weight between two users (x, v) as follow.

$$sim(x,v) = \frac{\sum_{i}^{N}(r_{x,i} - \bar{r}_x)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i}^{N}(r_{x,i} - \bar{r}_x)^2}\sqrt{\sum_{i}^{N}(r_{v,i} - \bar{r}_v)^2}}$$

Where N ($N = object_x \cap object_v$) represents the objects rated by both x and v, $r_{x,i}$ is the set of objects rated by user x and $\bar{r}_x$ is the average rating of user x

Predicting missing data. Collaborative filtering suffers a problem when one or more users did not want or ignored to evaluate the object, here we must predict their rating score, so after similarity computing, the system form a neighborhood N for each user and predict the rating of user U for learning object being created by computing the weighted average of the neighborhood users' rating using similarity calculate above as weights:

$$P = \bar{r} + \frac{\sum_w sim(u,u')(r_w - \bar{r}_w)}{\sum_w sim(u,u')}$$

Our work is destined for authors in order to help them achieve their goal in education content creation task, this system can be useful for novice authors which is strongly supported in our system. However the competencies of new authors are unknown for our database (situation known as cold-star in many filtering systems).

The problem of cold-start consist essentially in the following: a) recommendation of existing objects for new users, b) recommendations of new objects for existing users c) recommendations of new objects for new users [20]. Many approaches attempt to overcome this problem, most of them try to propose items to users in order to rate it at the beginning of their profile building or using stereotypes and/or asking users to answer questions related to their preferences. in our context which we considered that new users come to our system in order to create new learning objects, we adopt the content information to deduce similarities from existing objects compared to new objects, however we seem that a efficient similar users' set can find it using keywords' list i.e. the documents list retrieved is used to give all users that rate or create previously the document list and sorting them, the system show also what users are in learning object content. Creating data provide a solid proxy for eliciting user competencies (rule6) but generally give a small set of users especially when we specify the domain field, so to solve this inquiry we use the rating data to extend the users list (rule7) because the fact that a high score might imply that the user has really used the object or, at least is comfortable with it [26].

More formally, the users list is:

$$N = \{A_c \cup A_r\}$$

Where A_c represents the users' set that created and A_r represents the users' set that rated one or more learning objects, this learning object must have at least one of keywords' list. This formula aims to retrieve all users who have participate by rating or creating one or more learning objects similar to learning object being created, this set of users is given by rule6 and rule7. However this formula can lead to a big list of users (database increased over the time), we use the formula below in order to limit the above list (top N users selection).

$$P_i = \sum_{k=1}^{M} \#(C_k) + \beta \sum_{k=1}^{M} \#(N_k)$$

Where C_k represent how many time the keyword k appears in documents created by user I and N_k represent also how many time the keyword k appears in documents rated by user I. the factor β is a constant that can be parameterized depending on the activity in the system for weighting the creation task opposite the rating task, his range is between (0,1). At the end of this step and after collect all user score (predicted and data value) the system calculi the average (which represent the final score for learning object) in order to update/create the user profile and/or notify the user to revise his/her learning object if the score given was less than a threshold adjusted by the active user.

$$Avg = final\ score_{LO} = \frac{\sum_i^N r_i}{N}$$

The new metadata elements proposed in this work are fulfilled automatically using predefined rules; the result to be stored represented respectively the suitable context information and unsuitable context information retrieved by rule3 and rule8 respectively. So after gathering data we apply the algorithm below in order to retrieve suitable and unsuitable context information which represented as a vector represents respectively the bandwidth, screen size and screen resolution relating to learning object.

Input *: dataset of suitable and unsuitable context*
Output *: suitable Context vector* ***and*** *unsuitable context vector*
Foreach *element in (suitable_Context)* ***do***
    If *suitable_Context [i]<=*
OneOf(*Unsuitable_context[i])* ***then***
        Clear (Unsuitable_context [i])
    Suiable_Context := min(Suitable_Context[i])

Foreach *elements in Unsuitable_Context* ***do***
    Unsuitable_Context:= max (Unsuitable_Context [i])

The next example shows the obtained suitable and unsuitable context data. The problem recognized in such situation is how to make decision for end user about context suitability which can take any value.

TABLE3: EXAMPLE OF EXTRACTED CONTEXT INFORMATION

	Score	Resolution (Mpixels)	Size (inch)	Bandwidth (kb/s)	Suitable
User1	Like	1,2	4	1,024	yes
User2	Like	0,8	3,5	7,168	yes
User3	Like	0,9	6	0,512	yes
User4	Not provided	2,1	5	0,128	no
User5	Like	2,2	19	0,064	yes
User6	Like	2,1	15	7,168	yes
User7	Not provided	1,2	3,5	2,048	no

After running the algorithm our system will get the suitable and unsuitable context information (C_s) and (C_{us}) respectively, this dataset is considered as training set used to generate decision model for any learning content request carry out by end-users (learners) taking into account their context (C_i), the code below shown the prediction task

Input : *CS ,CUS ,Ci*
Output: *Suitability or Unsuitability of Ci*
If *((Ci[K] > CS[k]) or ((Ci[k] < CUS[k]))* ***then***
Begin
If *(Ci[k] > CS[k])* ***then***
    the user context is suitable
If (Ci[k] < CUS[k])) then
    the user context is not suitable
End
Else
  perform_suitability (Ci);

Where K denote the context type (resolution, screen size, bandwidth) and perform_suitability is a function that has one parameter represents the context data of the end user and returns the probability of C_i belongs to specific class (suitable or unsuitable). In this paper we adopt for baysian method to estimate the likelihood of specified context value is belongs to the suitable class or not. The Naïve Bayesian is powerful algorithms that provides high precision and speed treatment in vast capacity data compared to that of neural network algorithms or decision trees [16] used for classification task.

Given X as vector data of learner context in order to be classified in its class (suitable or unsuitable), and Y can be supposed that X is integrated in a class of C. The probability in which Y will happen as instance data of X is generated can be calculated as P(Y|X) which represents the prior probability.
The formula below is used to calculate P(Y|X).

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

Where
P(Y|X) is the posterior likelihood of class (Y) given predictor (X).
P(Y) is the prior likelihood of class.
P(X|Y) is the probability which is the probability of predictor given class.
P(X) is the prior likelihood of predictor.
Because our training data contains a continuous attribute x_i the probability distribution of x_i given a class C, p(X= x_i |C), can be computed by plugging x_i into the equation for a Normal distribution (Gaussian) parameterized by the mean μ and standard deviation σ. That is,

$$p(x_i|y) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x_i - \mu)^2}{2\sigma^2})$$

Where $\mu = \frac{\sum_{i=1}^{n} x_i}{n}$ and $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i\text{-}\mu)^2$

To determine the class of the target item, the following formula is calculated

$$P(x|C_i) = P(x_1, x_2, \dots, x_n|C_i)P(C_i)$$
$$= P(x_1|C_i)P(x_2|C_i) \dots P(x_n|C_i)P(C_i)$$

$$= P(C_i)\prod_{k=1}^{n} P(x_k|C_i)$$

The class that produces the highest or maximum probability is the classification for input data

$$C = argmaxP(C_i)\prod_{k=1}^{n} P(x_k|C_i)$$

And the prior probability $P(C_i)$ for each main category (suitable and not suitable) is 1/2 (as there are 2 categories)

System implementation and experiments

We have developed a tool for learning object creation task; it consists of a set of features provided to help authors to know the reliability of their educational materials, the user of our system must be registered or login through an interface provided by the system, in case of new user the system shows an additional form contains all user information's that needed by our system.

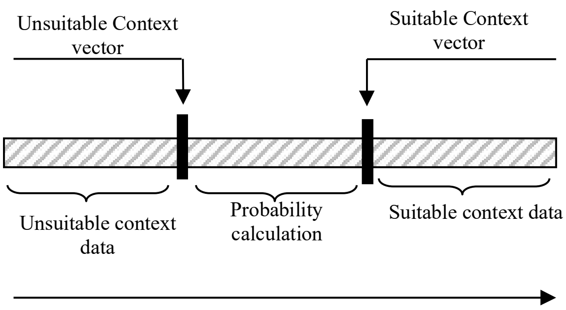


Figure 4: Vector of Context data Score

After that, the system shows a notification when the registered user was requested to evaluate another learning object or the user can begin create a new learning object or consulting the score of her/his earlier learning object. in this work we are implemented a server based system over internet where the server consists of database that stores the learning content, metadata, rating data and users profiles and the client side provides functionality for the establishment of the learning objects creation and rating task in the case of last one the system stores the contextual metadata like: screen size, resolution, internet bandwidth, the rating data and the contextual metadata are uploaded to a remote application server.

<technical>
<Requirement>
<Suitable_context name=”screen size” >
<value unit=”inch”> 7 < /value >
</ Suitable_context>
<Suitable_context name=”Resolution” >
<value unit=”Mpixels”> 0.8 < /value >
</ Suitable_context>
<Suitable_context name=”bandwidth” >
<value unit=”Mbps”> 0.512 < /value >
</ Suitable_context>
<Unsuitable_Context name=”screen size” >
<value unit=”inch”> 4 < /value >
</Unsuitable_ Context >
<Unsuitable_ Context name=”Resolution” >
<value unit=”Mpixels”> 0.5 < /value >
</Unsuitable_ Context >
<Unsuitable_ Context name=”bandwidth” >
<value unit=”Mbps”> 0.128 < /value >
</Unsuitable_ Context >
</Requirement>
</technical>

As experiments phase, our work is composed of two parts the first one based on the collaborative filtering in order to get a final score allows us improving the learning content and the second part is the extraction of context information to have dealing the outputs to end user taking into account his context. For the first part we use the recall, F-measure, and precision to evaluate the accuracy metrics of recommendation algorithm

In fact, the outputs of our recommendation algorithm contain two sets of users named positive participants and relevant participants, the positive participants are the users retrieved by our algorithm that rated the learning content and the relevant participants set which is the set of users who have been retrieved by our algorithm and not provide their rate, this set is devised on two subset negative relevant participants and negative relevant participants caused by their context (inappropriate context). To determine the accuracy metrics we put Np the set of positive participants which can seen as true result of the outputs of recommendation algorithm and Nr the set of relevant participants which can seen as true negative outputs

$$precision = \frac{Np}{Np + Nr}$$

$$recall = \frac{Np}{Np + (Nr - Nc)}$$

$$Fmesure = 2 * \frac{precision * recall}{precision + recall}$$

Where Nc represents the number of users which have an inappropriate context counted by our context extraction algorithm.

The purpose of second part is to make decision that a specific configuration represented as a vector (C_i ,C_j ,…) is suitable or unsuitable to use the learning object for this second part we report the performance evaluation result of the proposed data extraction method using empirical user study approach. we perform a sequence of test on platform of our university which we integrated our database on the server web application and the web application is distributed over many devices, we have supplied the basis to start this test with 65 users (teachers) and 18 learning objects of various form (text, multimedia, ...) on one single topic. The following is an extract that shows 14 participants whose 7 users have given their rate and 7 users does not provide their rates which require us to estimate their rating score.

Table 4: Example of training data

	Score	Resolution (M-p)	Size (inch)	B.W (Mb/s)	suitable
U1	Like	1,2	4,6	1,024	yes
U2	Like	0,8	4,5	2,048	yes
U3	Like	0,9	6	0,512	yes
U4	Like*	0,4	3,8	0,128	No
U5	Like	2,2	5	0,64	yes
U6	Like	2,1	5	1,048	yes
U7	Like*	0,5	3,5	0,128	No
U8	Like	1,3	4,4	1,024	yes
U9	Like*	0,6	3,3	0,056	No
U10	Like*	0,5	3,5	0,128	No
U11	Like	1,2	5	1,024	yes
U12	Like*	0,6	3,5	0,128	No
U13	Like*	0,4	3,8	0,056	No
U14	Like*	0,4	4	0,256	No

Where (*) denote predicted score. The "like" user attention implied that the user has given a high score for learning object After applying the extraction algorithm we obtain as suitable context data the vector (0.8 , 4.4 , 0.512) and unsuitable context data the vector (0.6 , 4 , 0.256) as shown in figure.

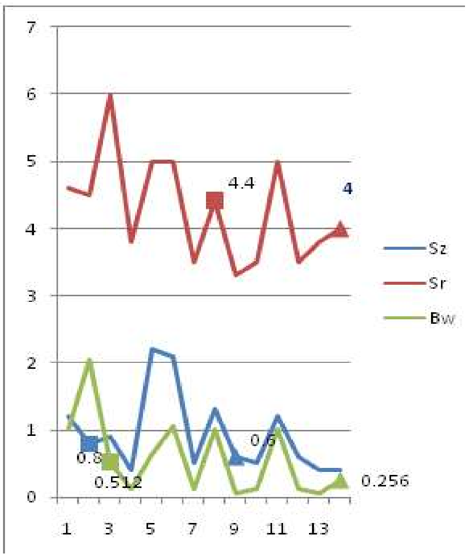
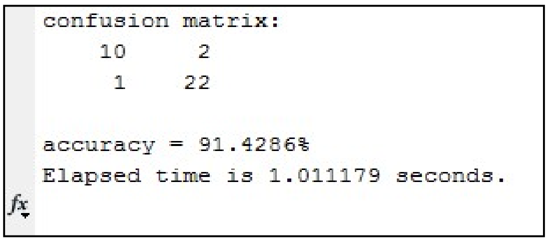


Figure 5: suitable and unsuitable values

In order to classify an input data for example (0.9, 3.2, 0.366) which represents respectively the screen resolution, screen size and bandwidth we calculate the probability using naïve bayes method with Gaussian distribution, for the above example we obtain P(yes)= 3.2391e-04 and P(no)= 1.7480e-07 which our system makes a decision that this configuration is suitable for using this leaning object. In order to identify common misclassifications we have calculate the confusion matrix [18] using Matlab framework, a confusion matrix contains information about actual and predicted classifications done by a classification system. A confusion matrix illustrates the accuracy of the solution to a classification problem. Our confusion matrix shows that the classification accuracy is very encouraging with minor errors as shown in figure below



```
confusion matrix:
    10     2
     1    22

accuracy = 91.4286%
Elapsed time is 1.011179 seconds.
```

Figure 6: Confusion matrix

5. Discussion & Conclusion

In this paper, we proposed a collaborative filtering based solution to improve the recommendation task by trying to detect the suitable and unsuitable context information concerning resources hardware information in order to deliver the education materials taking into account the context information of the end user, in this work we investigate the application of semantic web technologies to the building user profile with focus on rating data and user attention, we assume in this study that the user context plays a very important role on rating task and to evaluate the proposed approach we developed an tool-based authoring environment, this system enables rating and creating (or editing) of learning content compliant to the user's knowledge of the subject domain, this learning object are gathered into repository with its metadata that available for further use. In general, we can state that the proposed method can substantially improve the recommendation process taking into account information of user context, these last one is gathered throw monitoring and analyzing of user behaviors, also we can state that our method remains generic which can be applied with other contextual information like location and time, the success of this approach is situated in user behavior analysis to retrieve required context information that can be used in recommendation process without having to be identified by the owner of the object, in other hand this approach presents some limitation to apply it with other contextual information which requires that contextual information studied have an impact on object which w are put some assumptions at the beginning of this work. As future work, we want to achieve out more experiments that use different user profiles and knowledge areas. We want also to study other contextual information like location and time and analysis their impact on recommendation process. In this way, we could do a further validation of the effectiveness of our work.

REFERENCES

[1] Adomavicius, G., &Tuzhilin, A. (2011). Context-aware recommender systems. In Recommender systems handbook (pp. 217-253).Springer US.

[2] Albanese, M., d'Acierno, A., Moscato, V., Persia, F., & Picariello, A. (2010, July). A ranking method for multimedia recommenders. In Proceedings of the ACM International Conference on Image and Video Retrieval (pp. 311-318). ACM.

[3] Almeida, A., Orduña, P., Castillejo, E., Lopez-de-Ipiña, D., & Sacristán, M. (2011). Imhotep: an approach to user and device conscious mobile applications. Personal and Ubiquitous Computing, 15(4), 419-429.

[4] Berri, J., Benlamri, R., & Atif, Y. (2006, July). Ontology-based framework for context-aware mobile learning. In Proceedings of the 2006 international conference on Wireless communications and mobile computing (pp. 1307-1310). ACM.

[5] Bikakis, A., Patkos, T., Antoniou, G., & Plexousakis, D. (2007, November). A survey of semantics-based approaches for context reasoning in ambient intelligence. In European Conference on Ambient Intelligence (pp. 14-23). Springer, Berlin, Heidelberg.

[6] Chan, T., Sharples, M., Vavoula, G., & Lonsdale, P. (2004). Educational metadata for mobile learning. In Wireless and Mobile Technologies in Education, 2004.Proceedings. The 2nd IEEE International Workshop on (pp. 197-198). IEEE.

[7] Chen, H., Perich, F., Finin, T., & Joshi, A. (2004, August). Soupa: Standard ontology for ubiquitous and pervasive applications. In Mobile and Ubiquitous Systems: Networking and Services, 2004. MOBIQUITOUS 2004. The First Annual International Conference on (pp. 258-267). IEEE.

[8] Cheng, B., Zhong, S., & Chen, J. (2017). Context Ontology-Based Reasoning Service for Multimedia Conferencing Process Intelligence. IEEE Transactions on Systems, Man, and Cybernetics: Systems.

[9] Daniel, L. (2005). Discovering knowledge in data: An introduction to data mining. Larose, p. cm, ISBN 0-471-66657-2 (cloth), QA76, 9, D343L38.

[10] Ekstrand, M. D., Riedl, J. T., &Konstan, J. A. (2011). Collaborative filtering recommender systems. Foundations and Trends in Human-Computer Interaction, 4(2), 81-173.

[11] Erik, D. Neil, S., Marc, V. Application Profiles for Learning, Proceedings of the Sixth International ICALT'06, 2006

[12] Golemati, M., Katifori, A., Vassilakis, C., Lepouras, G., &Halatsis, C. (2007, April). Creating an ontology for the user profile: Method and applications. In Proceedings of the first RCIS conference (No. 2007, pp. 407-412).

[13] Gruber, T. R. (1993). A translation approach to portable ontology specifications. Knowledge acquisition, 5(2), 199-220.

[14] Hong, K., Jeon, H., & Jeon, C. (2013). Personalized Research Paper Recommendation System using Keyword Extraction Based on UserProfile.Journal of Convergence Information Technology, 8(16), 106.

[15] IEEE Learning Technology Standards Committee. (2002). Draft Standard for Learning Object Metadata (LOM) IEEE 1484.12. 1.

[16] Kim, J., Lee, D., & Chung, K. Y. (2014). Item recommendation based on context-aware model for personalized u-healthcare service.Multimedia Tools and Applications, 71(2), 855-872.

[17] Koren, Y., & Bell, R. (2011). Advances in collaborative filtering. In Recommender systems handbook (pp. 145-186).Springer US.

[18] Margaret H.. Dunham. (2003). Data mining: introductory and advanced topics. Prentice Hall.

[19] Nilsson, M., Palmér, M., & Brase, J. (2003). The LOM RDF binding: principles and implementation. In Third Annual ARIADNE conference, Leuven Belgium, 2003.

[20] Park, S. T., & Chu, W. (2009, October). Pair wise preference regression for cold-start recommendation. In Proceedings of the third ACM conference on Recommender systems (pp. 21-28).ACM.

[21] Ramos, J. (2003, December). Using tf-idf to determine word relevance in document queries. In Proceedings of the first instructional conference on machine learning.

[22] Razmerita, L. (2011). An ontology-based framework for modeling user behavior—A case study in knowledge management. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 41(4), 772-783..

[23] Rodríguez, N. D., Cuéllar, M. P., Lilius, J., & Calvo-Flores, M. D. (2014). A survey on ontologies for human behavior recognition. ACM Computing Surveys (CSUR), 46(4), 43.

[24] Santos, O. C., Granado, J., Raffenne, E., & Boticario, J. G. (2008, November). Offering recommendations in Open-ACS/dotLRN. In Int. Conf. on Community based environments, Valencia.

[25] Schirru, R., Baumann, S., Memmel, M., & Dengel, A. (2010). Extraction of Contextualized User Interest Profiles in Social Sharing Platforms. J. UCS, 16(16), 2196-2213.

[26] Sergis, S., Sampson, D. (2016). Learning object recommendations for teachers based on elicited ICT competence profiles. IEEE Transaction on Leaning Technologies, 9(1), 67-80

[27] Sergis, S., Zervas, P., & Sampson, D. G. (2014, July). Towards learning object recommendations based on teachers' ICT Competence Profiles. In 2014 IEEE 14th International Conference on Advanced Learning Technologies (pp. 534-538). IEEE.

[28] Tsai, K. H., Chiu, T. K., Lee, M. C., & Wang, T. I. (2006, July). A learning objects recommendation model based on the preference and ontological approaches.In Advanced Learning Technologies, 2006. Sixth International Conference on (pp. 36-40). IEEE.

[29] Vanderdonckt, J., Grolaux, D., Van Roy, P., Limbourg, Q.,

Macq, B., Michel, B.: A design space for context-sensitive user interfaces. In: Proceedings of IASSE (2005)
[30] Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., & Duval, E. (2012). Context-aware recommender systems for learning: a survey and future challenges. IEEE Transactions on Learning Technologies, 5(4), 318-335.
[31] Wang, X. H., Zhang, D. Q., Gu, T., & Pung, H. K. (2004, March). Ontology based context modeling and reasoning using OWL. In Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on (pp. 18-22). Ieee.
[32] Zhang, D., Gu, T., & Wang, X. (2005). Enabling context-aware smart home with semantic technology. In International Journal of Human-friendly Welfare Robotic Systems
[33] G. Klyne, F. Reynolds, C.Woodrow, H. Ohto, J. Hjelm,M. H. Butler, and L. Tran. 2004. Composite Capability/Preference Profiles (CC/PP): Structure and Vocabularies 1.0. W3C Recommendation. Retrieved from http://www.w3.org/TR/CCPP-struct-vocab/.