Ontology Based e-Learning Systems: A Step towards Adaptive Content Recommendation

S. Sarwar, R. García-Castro, Z. Qayyum, M. Safyan, F. Munir, and Muddesar Iqbal

Abstract—E-Learning systems can be made more effective through personalization and adaptivity while recommending the learning content to learners. A comprehensive set of attributes needs to be identified for learner categorization to ensure personalized and adaptive content recommendation. In this paper, a set of core attributes have been identified for effectively profiling the learners and categorizing through neural networks. The learning contents have been annotated formally in ontology for recommending the personalized contents to the learners. Performance of proposed framework is measured in terms of accurate learner categorization, precise recommendation of the learning contents and completeness of ontological model.

Index Terms—Adaptivity, content recommender e-learning, personalization.

I. INTRODUCTION

The phenomenon of e-learning has prevailed a lot, especially in last decade, entailing in removal of numerous educational barriers such as information access, cost (of traveling, class room environment, teaching infrastructure), location and time in educational as well as corporate sectors. Apart from benefit of indiscriminant global reach (a.k.a. universality) to information repositories, e-learning provides up-to-date information ensuring high productivity and relevance to dependent as well as independent learners. e-learning is not merely confined to prompt deliverance of educational contents, rather it’s a line of packages ranging from content development to maintaining profile of learners, aligning contents to respective learners as per their ability, from maintaining practice exercises to managing grading, adaptivity and personalization of learning material and searching from relevant educational repositories.

With core functional components and services [1], focus of our work remains on learner specific contents deliverance (personalization), sequencing/re-sequencing of learning contents based on learner abilities, recommendation of contents to the learners and compliance of these learning artifacts to context-aware web 3.0. Each of these aspects is highlighted in the following.

Learning Attributes Profiling for Personalization of learning contents ensures that only relevant contents are presented to learner with respect to his cognitive abilities. It is contrary to a typical “one size fits all” approach that may not fully comprehend the learner’s capacity to learn while presenting learning contents. Different attributes of learners have been considered while offering the learning contents instead of presenting same contents to all the learners. Once profiles are developed based upon these attributes, learners are categorized into different categories for ensuring personalization as shown in [2]-[7]. The proposed technique exploiting the Artificial Neural Networks (ANN) for learner categorization that uses learner profiles (combination of learner attributes) dynamically. Learners are categorized into one of the four categories of 'Novice', 'Easy', 'Proficient' or 'Expert' based on their profiles. Here it is worth mentioning that these learner categories were devised after a survey from the educational psychologists, learner evaluation from the behavioral and cognitive perspectives, literature [3], [4], [6], [7] and input from seasoned educationists and academicians.

Learning Content: Web technologies have a paramount role in ongoing success of e-Learning applications for incorporating usability of learning contents, design of learning components and availability of digital repositories [8], [9]. However, everything on current version of web is syntactical, machine-readable and not machine-understandable [10] making e-Learning solutions less flexible and less interoperable. Furthermore, context-aware alignment of contents to learners initially and after assessment is also an open research issue. These problems suggest porting of learning contents to context aware Semantic web (or Web 3.0) based on ontologies. Once learning content has been established in the form of ontologies and learner categories [11] have been assigned to the learners, mechanism to recommend the learning contents needs to be devised for effective and efficient learning of the learners.

Learning Content Recommender: In order to make recommendations of learning content for learners, collaborative or content based educational recommenders use the ratings corresponding to learners, learning contents and content features [12], [13]. However, they may have issues of cold start/ramp-up (lack of historic information), early rater (no rating Information), and overspecialization (lack of knowledge about learner’s level) suggesting knowledge base recommender as a solution.

Keeping above in view, a rule based knowledge driven
recommendation is presented that models the profiles of the learners and the learning contents in ontologies. Learning contents are recommended to the learners based upon their profile categories contained in ontology (No dependence on Ratings due to machine learning techniques based learner categories). Moreover, its helps in assessment based content alignment i.e. adaptivity [8] of learning content. Detail of this recommender is furnished in section 3.

Rest of the paper is organized as follows: Section II provides a brief review of web 3.0 compliant e-learning systems for effective learning. Section III presents the proposed model of e-learning framework in a layered view along with details of approach followed, Section IV provides an overview to implementation strategies of framework and Section V concludes the work with a view of future directions.

II. LITERATURE SURVEY

A handful of literature is available on different aspects of e-learning such as rationale for e-learning system, modeling the learner profiles and learning contents for subsequent categorization, aligning learning contents and learner model with semantic web technologies and recommendation of learning contents to the learner etc. Each of these aspects is briefly reviewed in the following:

This paper [14] works on the philosophy that learner’s style should not be focused than learner’s ability for personalization. Tests are used to estimate learner’s ability dynamically. Different models such as domain model (classes/properties describing topics of domain and pedagogical relations), learner model (for learner’s profile, preference and identification) and content models have been developed for building respective ontologies. Lastly, adaptive engine generates personalized contents based on learner’s information coming from learner’s model.

Continuity in number of exercises recommended to learner is a major issue. Several aspects of “Web usage mining” have been presented but current solution employs ontologies (ontology matching), fuzzy logic (fuzzy classifiers) and tree hierarchy (namely M-Tree) [5]. Also, target is to make system more adaptive, semantically rich and swift to respond the learner queries. Given a query, it is subjected to TD-IDF (Term Frequency; Inverse Document Frequency) followed by ontology reasoning and fuzzy logic. Lastly, collaborative recommendation of contents is done via Pearson’s correlation. Mean Absolute Error (MAE) and Normalized Mean Absolute Error (NMAE) is used to measure the recommendation quality of recommendations. Smaller value of NMAE implies a better quality.

An educational tool (named LOD-Learning) based on semantic technologies is introduced for enhancing the contents of learners in given subject [15]. Widespread implementation of e-learning systems has introduced Learning Management Systems (LMS) for sharing, creating and collaborating the learning contents. However, these contents are context-unaware. Therefore, a semantic technology esp. RDF is used for learning course content’s expression [16]. Here enhancement refers to adding/linking related material/content to the current lessons. Most relevant contents are searched using NLP entity recognition algorithm for extracting and incorporating out of given text. LOD-Learning content is compiled with Sekai LMS that in turn uses IMS learning standard protocol along DBpedia. Effectiveness of technique is evaluated in terms of learner’s performance in an inter-subject study in a defined time slot by comparing “Before-Tool” and “After-Tool” effect-size. (empirical research is still in progress).

In [17], focus of work is to employ semantic technologies for persistence services in Learning Management Systems (LMS) through more expressive/flexible/heterogeneous/reusable representations. “Online communities” connected the external content base with learners/teachers to develop the “virtual communities”. Also, internal and external contents have been linked to provide graph-based navigation in platform. A step ahead is SIOC (Semantically Interlinked Online Communities) that represents rich data from social web. Idea presented was applied on legacy applications by ending up in RDF graph that can be queried with SPARQL. Final framework was tested on big-data based technologies like Hadoop, Hbase, Flink etc. The entity centric nature of technique helps in identifying the resources via URI while competing the “linked data” environment for e-learning resources in communities.

A system [18] for betterment of knowledge management and for representing the associated data in learning management systems contrary to previous systems. A domain ontology along with profile ontology has been presented through VARK model for learner’s classification. A comprehensive view of m-learning is given along with semantic technology for location-independent learning. The aspect of personalization for learning content provision keeping in view capacity and skills of learner have been provided through VARK model of learning. ACM computation classification has been used to get baseline concepts for domain ontology. Profile of learner has been built by acquiring demographic information of the learner. Organization of the learning contents offered to the learner has not been elaborated though it is the core contribution of work as claimed. Also the aspect of personalization to recommend the relevant contents seems missing. Development of feedback system is envisioned as a future work. Also, it may aid in taking a step towards IoT (Internet of Things) based e-learning system.

Lastly, few techniques [19], [20] claim to target the semantic web but formal and explicit descriptions of learners and learning contents using ontologies seem missing.

III. PROPOSED ARCHITECTURE

On headway to developing the proposed semantics based e-learning recommender, ultimate goal was to provide personalized and adaptive learning contents to the learners. Layered architecture of proposed system, as shown in Fig. 1, provides an insight into the proposed architecture with respect to how it has been embedded in current technological layers and development strategies used for effective delivery of learning contents to facilitate learners.

There are three main layers of proposed framework
namely (1) Modeling Layer, (2) Core Functional Layer and (3) Persistent Layer. Each of these layers is further segregated into “work packages” facilitating in division, management and maintenance of learning cycles. Each of these layers is elaborated in the following:

**Modeling Layer:** This layer provides a point of interaction to the proposed system for learner as well as instructors in order to perform different tasks such as course registration, viewing of contents, and course administration. From technology perspective, interface comprises of a combination of HTML/JSP pages interacting with core functional layer through Java Servlets which in turn performs data management operations with persistent layer.

**Core Functional Layer:** This layer covers main implementation portion of proposed system from perspective of semantic web technologies. Core functional layer has been further divided into following three functional sub-layer with respect to their operational profile:

Primary Functions Sub-layer: This layer is responsible for management of learner characteristics/profiles and learning contents termed as Sharable Content Objects (SCOs). Score, Feedback, Log: This working unit manages the feedback mechanism based on score of learners and logs the grades/score for content sequencing. Logging of these aspects enabled us to incorporate personalization in learning contents for the learners (as explained in section 2.5). SCO Searching provides the facility of context-aware searching to the learner. Initially search based on topics/ sub-topics has been incorporated.

Secondary Functions Sub-layer: This layer deals with management of some pre-requisites for working with proposed system e.g. user authentication, authorization etc. The foremost step for any learner is to register by giving the credentials so that system can authenticate the learner(s) for subsequent interactions. Other activities, specific to this layer to be catered in future, are announcements, tutoring and administration of overall system (part of future goals).

Tertiary Functions Sub-layer: This layer comprises of non-functional requirements but have a key role in making the system effective through aspects of personalization and adaptivity (both aspects are discussed in detail in section 2.5). Adaptivity and personalization are carried out through categorization of leaner profile concepts in “LearnerOntology”, performance measured through “AssessmentOntology” and in coordination with “CourseOntology” through a semantic recommender system as given in section 3.3. Moreover, these functions are responsiveness and delivery of contents efficiently in a timely manner. This aspect is dependent on efficient management and retrieval of contents from persistent layer.

A detailed view of learner attributes with an effective role in categorizing the learners, as exhibited during experimentation, needs to be discussed before getting into the details of categorizing the learner and respective content recommendation. Also, the way data was acquired to build the learner’s profile and for recommending the learning contents has been presented as follows:

**Learning Aptitude Test:** provides a basis for predicting an individual’s ability, with training, to acquire some knowledge, skill, or set of responses. Also, it predicts an individual’s potential with aptitude test scores. There are different domains covered in every test such as numerical aptitude, analytical skills, mechanical reasoning and verbal reasoning. Every test has certain score in ratio of total score but it does not depict the cognitive level of individual. Some meanings have to be attached to the score that can be in the form of Percentile or Stanine [3] as shown in Table I.

**TABLE I: STANINE SCALE FOR HAVING LEARNER’S APTITUDE**

<table>
<thead>
<tr>
<th>Percentile</th>
<th>V. Low</th>
<th>Below Avg.</th>
<th>Average</th>
<th>Above Avg.</th>
<th>V. High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>0.0</td>
<td>1-2</td>
<td>3-4</td>
<td>5-6</td>
<td>7-8</td>
</tr>
<tr>
<td>5-6</td>
<td>7-8</td>
<td>9-10</td>
<td>11-12</td>
<td>13-14</td>
<td>15-16</td>
</tr>
<tr>
<td>7-8</td>
<td>15-16</td>
<td>17-18</td>
<td>19-20</td>
<td>21-22</td>
<td>23-24</td>
</tr>
<tr>
<td>8-9</td>
<td>23-24</td>
<td>25-26</td>
<td>27-28</td>
<td>29-30</td>
<td>31-32</td>
</tr>
<tr>
<td>9-10</td>
<td>31-32</td>
<td>33-34</td>
<td>35-36</td>
<td>37-38</td>
<td>39-40</td>
</tr>
<tr>
<td>10-11</td>
<td>39-40</td>
<td>41-42</td>
<td>43-44</td>
<td>45-46</td>
<td>47-48</td>
</tr>
<tr>
<td>11-12</td>
<td>47-48</td>
<td>49-50</td>
<td>51-52</td>
<td>53-54</td>
<td>55-56</td>
</tr>
<tr>
<td>12-13</td>
<td>55-56</td>
<td>57-58</td>
<td>59-60</td>
<td>61-62</td>
<td>63-64</td>
</tr>
<tr>
<td>13-14</td>
<td>63-64</td>
<td>65-66</td>
<td>67-68</td>
<td>69-70</td>
<td>71-72</td>
</tr>
<tr>
<td>14-15</td>
<td>71-72</td>
<td>73-74</td>
<td>75-76</td>
<td>77-78</td>
<td>79-80</td>
</tr>
<tr>
<td>15-16</td>
<td>79-80</td>
<td>81-82</td>
<td>83-84</td>
<td>85-86</td>
<td>87-88</td>
</tr>
<tr>
<td>16-17</td>
<td>87-88</td>
<td>89-90</td>
<td>91-92</td>
<td>93-94</td>
<td>95-96</td>
</tr>
<tr>
<td>17-18</td>
<td>95-96</td>
<td>97-98</td>
<td>99-100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PreTests were developed for students with varying level of difficulty keeping in view their prior knowledge and pre-requisite score while consulting with domain experts. Pre-Test in our case comprised of programming questions developed using GAT Subject (CS) test banks [21], exercises from text books [22] and online quizzes [23]. Stanine standard used to categorize the learners based upon scores obtained.

Some other attributes are the qualification, professional experience, age, pre-req GPA and CGPA. The Learner Attributes acquired from different sources have been used to categorize the learner profiles with respect to their level of expertise and cognitive skills [24]. Few of the restrictions on ontology concepts through Description Logic (DL) constructs are illustrated in Fig. 2. Such representation of learner’s profile as an ontological knowledge-base can provide an optimal representation of information with consistency among concepts, reduced redundancy and capacity to infer and reason for intelligent information retrieval. The process of acquiring the learner profile attributes from implicit as well as explicit sources has been explained in section III.A.

This repository of learners has been used by Neural Networks for categorization of learners. The performance of machine learning techniques greatly relies on the quality of the data set used for training, so it is important to provide a glimpse of such dataset. In total, there were profiles of 600 learners, each having 12 profile attributes for correctly categorizing the learners.

**A. Learner Categorizer**

The Multilayer Perceptron (MLP) model of Artificial
Neural Networks (ANN) has been employed for learner categorization. These learner categories have been used to provide the learner with suitable learning content. The MLP has been selected due to its ability of regulating network weight in order to minimize the Mean Square Error (MSE). The MLP model was implemented using Neural Pattern Recognition tool of Matlab 2015a with standard weights and activation functions. Besides, another script was written in Matlab separately for experimenting with different number of neurons and middle layers. The input layer contained 7 neurons, 2 hidden layers each with 8 neurons and an output layer with 1 neuron was used.

The performance of the model during validation/testing phases has been measured in terms of how accurately learners have been classified while considering the associated costs:

**B. Learning Content Model**

Learning contents, represented by topics and subtopics each corresponding to a concept in ontology as shown in Fig. 3, are also categorized based upon level of difficulty, programming language (since three languages are modeled in ontology), and number of weeks. Mapping among respective concepts from learner ontology and learning contents ontology is provided through a rule based system. On one hand, rule based recommender ensures aspect of personalization and on other side performance (or evaluation) based adaptivity is incorporated.

**C. Knowledge Based Recommender**

Once contents of a topic are modeled in ontology and connected to respective LOs, there is need to recommend each of the learners with learning content by Knowledge based Adaptive Semantic E-Learning Recommender (KASER). For example, a novice level learner is recommended with learning contents of novice type through personalization. Learner’s performance is tracked over the weeks for sequencing or re-sequencing the learning contents by KASER based on relative grading system.

**IV. RESULTS AND EVALUATION**

The proposed system has two aspects for evaluating the Ontology models through ontology evaluation techniques and evaluation to measure effectiveness of proposed techniques in terms of learner’s performance. Later evaluation of technique has two aspects (a) Accuracy of learner categorization (b) Precision of KASER in right content recommendation.

First, evaluating the health of ontologies that have been used in the content model i.e. Ontoclean [25] and OOPS (Ontology Pitfall Scanner) [26].

OntoClean meta-properties have been applied on classes and properties to check for subsumption relations and ontology model with violations. The patterns with violations in model were detected and removed using SPARQL queries. Fig. 4 asserts the fact of OntoClean which offers logic based argument in validating the ontological and taxonomical relationships through meta-properties of Rigid, Identity and Unity.

A technique named OOPS (Ontology Pitfall Scanner) has been employed to check if there is any fault that not aligned with quality standards. These categories are normal, minor, important and critical. OOPS findings for LearnerOntology are given in Table II with no critical pitfall.

**TABLE II: OOPS BASED EVALUATION OF LEARNER ONTOLOGY**

<table>
<thead>
<tr>
<th>Pitfalls</th>
<th>Ontology Pitfall</th>
<th>Inappropriateness Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural</td>
<td>Modeling Decision - P[24]</td>
<td>Normal</td>
</tr>
<tr>
<td>Functional</td>
<td>Real World Modeling - P[4,12]</td>
<td>Important</td>
</tr>
<tr>
<td>Functional</td>
<td>Application Context P[14,37,39]</td>
<td>Critical</td>
</tr>
<tr>
<td>Profiling</td>
<td>Ontology Clarity - P[9,22]</td>
<td>Normal</td>
</tr>
<tr>
<td>Consistency</td>
<td>P[3]</td>
<td>Important</td>
</tr>
<tr>
<td>Completeness</td>
<td>P[4,10,11,13]</td>
<td>Normal</td>
</tr>
</tbody>
</table>

Fig. 4. OntoClean based evaluation of learner ontology.
A class of 40 learners was selected for evaluation purpose with profiles available in learner ontology. These profiles (annotated in ontology with different attributes) were used for recommending the suitable contents. In order to compare and analyze the accuracy of recommendations made by the proposed system named KASER in view the profiles of the learners.

Contents recommended by learner were also recommended by domain experts in order to assert the degree of correctness shown by proposed system. In order to assert the level of agreement between contents recommended by KASER and the ones recommended by Domain Expert (DE), Kappa coefficient has been used as shown in Table III.

An average of contents recommended by domain experts was taken. This average was used alongside the contents recommended by KASER for calculating the Kappa’s coefficient. The range of Kappa’s co-efficient is shown from 70 to 85. As per research standards anything above 65% is adequate level of agreement.

TABLE III: KASER’S PERSONALIZED CONTENT RECOMMENDATION VALIDATED BY DOMAIN EXPERTS (DE)

<table>
<thead>
<tr>
<th>No. of Weeks</th>
<th>Learner’s Weekly Performance</th>
<th>Conventional Learning</th>
<th>KASER</th>
<th>A-KASER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

Fig. 5 provides a detailed insight to the weekly progress that learners made while following the learner activity through conventional system, with semantic recommender (KASER) and through adaptive semantic recommender (A-KASER). The assessments were designed such that depth of knowledge, coverage of contents and impact of remedial exercises could be evaluated.

The average of scores that learners acquired in weekly assessments, exercises and quizzes is taken for all 8 weeks of summer semester in 2015 for course of “Object Oriented Programming”. Here, assessment results of learners were recorded without taking into account any of the processes in proposed approach.

Fig. 5. Learner’s performance: Weekly impact of KASER and A-KASER.

V. CONCLUSION AND FUTURE DIRECTIONS

A component of our Ontology based e-learning system has been presented that proposes to categorize the learners using neural networks keeping in view the profiles of the learners. The learning contents which have been modeled in the form of ontologies are recommended to the respective learners with respect to their level of difficulty.

We look forward to employ the Random Forest algorithms [27] for categorizing the learners and respective content recommendation for. Moreover, an end-to-end ontology based system will be developed aligned with implementation perspective as discussed.

ACKNOWLEDGMENT

Authors of this manuscript are thankful for the support of the Ontology Engineering Group (OEG) in unconditional support and guidance for completing this research. Also, we wish to acknowledge our gratitude to the Higher Education Commission of Pakistan in extending its resources in completing this work.

REFERENCES


Sohail Sarwar received the M.S. degrees in information technology from National University of Science and Technology, Islamabad, Pakistan. He is currently pursuing the Ph.D. degree in the Department of Computing, University of Gujrat Pakistan. His research interests include e-learning, semantic technologies and knowledge engineering techniques. Lately, he has been part of Ontology Engineering Group, Universidad Politécnica de Madrid, Spain for research visit.

Raúl García-Castro is assistant professor at Universidad Politécnica de Madrid, Spain. He got his PhD degree in computer science and artificial intelligence from Universidad Politécnica de Madrid, Spain in 2008. His research interests include semantic technologies, ontology engineering, and ontology-based data integration.

Zia Ul Qayyum is currently a professor at University of Gujrat Pakistan. He received his Ph.D. degree in computer science from Leeds University UK in 2005. His research interests include artificial intelligence, knowledge engineering, data mining, semantic web and e-learning.

Muhammad Safyan is in Government College University (GCU) Lahore. He received his MS degree from National University of Sciences and Technology in 2009. His area of interest is ontology alignment, e-learning and semantic activity recognition.

Faisal Munir is research scholar in University of Catalan Spain and working on domain of quality control for ontology based system in perspective of big data and e-learning systems.

Muaddasser Iqbal is working as senior lecturer in London South Bank University England. He has been active part of the pool of our domain experts for evaluating the proposed model.