Ontology based Learning Framework: A Personalized, Adaptive and Context Aware Model

Sohail Sarwar1, Zia Ul Qayyum1, M Safyan1, Muddessar Iqbal, Anwer Al-Dulaimi, Raul García-Castro2

**Abstract**— Enhancing the degree of learner productivity, one of the major challenges in learning systems, may be catered through effective personalization, adaptivity and context awareness while recommending the learning contents to the learners. In this paper, a learning framework has been proposed that profiles the learners, categorizes the learners based on profiles, makes personalized content recommendations and performs assessment based content adaptation. A mathematical model has been proposed for learner categorization using machine learning techniques (a hybrid of case based reasoning and neural networks). The learning contents have been annotated through *CourseOntology* in which three academic courses (each for language of C++, C# and JAVA) have been modeled for the learners. A dynamic rule based recommender has been presented targeting a ‘relative grading system’ for maximizing the learner’s productivity. Performance of proposed framework has been measured in terms of accurate learner categorization, personalized recommendation of the learning contents, completeness and correctness of ontological model and overall performance improvement of learners in academic sessions of 2015, 2016 and 2017. The comparative analysis of proposed framework exhibits visibly improved results compared to prevalent approaches. These improvements are signified to the comprehensive attribute selection in learner profiling, dynamic techniques for learner categorization and effective content recommendation while ensuring personalization and adaptivity.

**Index Terms**— Ontologies, Learning, Personalization, Adaptivity, Content Recommender

—————————— ◆ ——————————

# 1 Introduction

T

he perceptible dominance of internet has affected every aspect of human life that can specifically be observed on academic landscape in the form of Electronic Learning (or “E-Learning). The revolutionary advent of knowledge universality through E-Learning has entailed in upsurge of educational elite in societies through development of professionals by eliminating costs of workforce and infrastructure. Currently, these systems are playing a pivotal role in transforming information societies into knowledge societies through widespread delivery of didactic contents with a vision to educate future generation of learners (expected to grow from 100 million learners to 250 million by 2025 [1]). Such cyberspace driven learning is paving the way for building hubs of inventive activities through emergence of E-Learning concepts [2] with target of maximizing the learner’s productivity and effectiveness.

E-Learning is not confined merely 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. The multifaceted and ubiquitous view of an E-Learning system has been illustrated in Fig. 1, aiming to improvise the role of E-Learning from information transmission to knowledge-construction [1] and deliverance targeted to enhance learner productivity.

E-Learning

Repositories

Curriculum Design

Personalization

Searching

Assessments

Academia

Adaptivity

Fig. 1: Dimensions of E-Learning System

With core functional components and services, focus of our work remains on deliverance of learner specific contents (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.

*Learner Attributes Profiling and 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. These attributes are learner’s academic performance, learning style, aptitude, background knowledge, and term-wise performance during the course etc (methods for acquiring these attributes are discussed in section 3). Based on these attributes, learners are categorized into certain categories for recommending the suitable learning content (sequenced with respect to difficulty levels).

*Learner Categorization:* Once the picture of the learner’s ability is clear by rightly categorizing the learner, learning contents may be offered keeping in view the learner’s ability. This categorization has potential to offer twofold benefits i.e. what contents to be offered to the learner for first time and adaptivity of contents (performance-based sequencing/re-sequencing of contents during the course). Different techniques have been proposed for learner categorization [2], [3], [4], [5], [6],[ 7] that categorize the learners. However, these techniques have issues such as: considering academic attributes subjectively, ignoring learners’ behavioral and demographic attributes, sort the learners into good and bad learners only, no reusing of information already categorized learners etc. The proposed technique exploiting the Artificial Neural Networks (ANN) for learner categorization 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. Once learners are categorized, a mechanism is desired to model the learning contents for every category of learners compliant with latest web technologies.

*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 [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. Proposed framework exploits all the benefits associated with Web 3.0 through different ontologies i.e. course ontology models course of “Object Oriented Programming”, assessment ontology to model quizzes and exercises. All the concepts (topic/sub-topic) in course ontology are annotated with a difficulty levels. These difficulty levels have been assigned to each topic or sub-topic with help of domain experts from industry as well as academia in compliance with prevalent standard such as SCORM/IMS [11]. Once learning content has been established in the form of ontologies and learner categories have been assigned to the learners, mechanism to recommend the learning contents needs to be devised for effective and efficient learning of the learners.

*Knowledge based 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). Knowledge base recommenders are gaining much attention in E-Learning systems for not depending upon different ratings as collaborative or content based recommenders do. Another aspect of knowledge base recommenders is the use of ontologies that are naturally right representation of learning objects mimicking their relationships and interdependence (in same way as a “concept map”). Moreover, ontologies are equipped with features of reusability, interoperability and share-ability across different platforms while preserving the semantics of learning contents.

Keeping above in view, a rule based knowledge driven recommender is presented that semantically 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). Detail of this recommender is furnished in section 3.

Rest of the paper is organized as follows: section 2 provides a brief review of web 3.0 compliant E-Learning systems for effective learning, section 3 presents the proposed model of E-Learning framework in a modular view along with details of approach followed, section 4 provides evaluation strategies and performance of framework and section 5 concludes the work with a view of future directions.

**2 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:

A philosophy, stating that learner’s style should not be focused than learner’s ability for personalization, has been discussed in [14]. 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.

An unsupervised classification technique, linear regression, is used for modeling the quantity of accumulated knowledge [5] pertinent to a learner. It uses variables linked to the learning activity, user experience and accumulated knowledge. Categorization of learners is performed at concept level that in turn is evaluated based on percent concepts covered in knowledge. After assessments, learners are classified based upon two aspects i.e. quality of answers and the time consumed in answering. Contents of the discipline were divided into chapters where every chapter was modeled into concepts at a concept map. A classified model was created based on aggregates as a result of learner’s experience in eventuation of learner’s experience for recommending chapters/concepts in a discipline. 500 students, 5 disciplines, 5-10 chapters, 10-20 exams/tests were used to initiate learning and concept coverage to evaluate the predictive accuracy of approach in recommending learning contents.

An E-Learning system [15] based on cross similar ontologies has been proposed for users whose profile is maintained after getting his evaluation score from the perspective of domain specific and domain independent aspects. Here, the learner’s profile (or learner model) for offering the learning contents is solely a function of assessing the score of the learner. Similarly, student’s personality is maintained conceptually through coupling knowledge management techniques and ontologies in [16] (empirical research is still in progress).

The focus of work in [17] 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, and 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 for betterment of knowledge management and for representing the associated data in learning management systems has been discussed in [18]. 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.

# 3 Proposed Framework

On headway to developing the proposed semantics based E-Learning system, ultimate goal is to provide personalized and

adaptive learning contents to the learners. This requires an

insight into the cognitive abilities of the learners that can possibly be acquired once learner’s attributes are profiled. So the proposed framework starts with accumulating the attributes of the learners as learner profiles followed by preprocessing of learner profiles to figure out most important ones having an impact on learner’s performance. Based upon attributes unleashed by preprocessing phase, learners were categorized into different categories aligned with their learning abilities. These learner categories were used for recommending the learning contents (annotated with difficulty level) in a personalized way followed by adaptive recommendation of learning contents. An architectural view of proposed system is illustrated in Fig. 2.

**LearnerProfile Preprocessor**

**LearnerProfile Categorizer**

**Personalized Recommender**

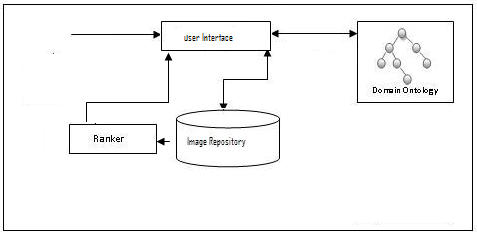
**Adaptive Recommendations**

**System Interface**

**Learner Attribute Profiler**

**Learner Attributes - Explicit**

**Learner Attributes- Implicit**



**Learning Content Model**

**Personalized/Improved Learning**

**Learning Cycle**

**Learning Cycle**

Fig 2: Architecture of Proposed System

**3.1 Learner Attributes Profiler**

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.

* Percentile is the percentage of candidates who fall below a particular raw score. e.g. a percentile of 65% means learner has better score than 65% of the class.
* Stanine [3] is the series of score represented on a scale from 0 to 9. Each of the score ranges classifies the given learner-score in to different categories. Stanine has been adopted as a scale to measure and categorize the learner based upon score achieved by each learner as shown in table 1.

Table 1: Stanine Scale for having Learner’s Aptitude

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Stanine** | **V. Low** | | | **Below Avg.** | **Average** | | **Above Avg.** | | **V. High** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Percentile | <= 4 | 5-11 | 12-22 | 23-40 | 41-58 | 59-77 | 78-89 | 90-95 | >97 |

*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.

*Qualification* of the learner is also a measure to decide upon cognitive skills of the learner i.e. to see if learner has attained undergrad, graduate or postgraduate courses while taking up the current course. Assumption: For current scenario, assumption is that learners have opted for undergraduate course (programming course).

*Professional Experience* is another factor in deciding upon how quickly, a learner can learn new concept/technology. There are three levels of experience considered in our case i.e. entry level, middle level and management level of experience.

*Learning Style* is a phenomenon to assess the suitable teaching method for a learner that suits the learner. This decision is based upon description of attitude and behaviour, learning skill/style can be measured. Felder-Silverman model [24] has been used to measure the learning style of learner as shown in Fig. 3. It represents four dimensions of learner’s personality having an impact on learning ability of leaner especially for hypermedia courseware. These dimensions are (1) Active/reflective (2) Sensing/intuitive (3) Visual/ verbal (4) Sequential/global. Learners are provided with a questionnaire to assess their learning behaviour for subsequent provision of learning contents.

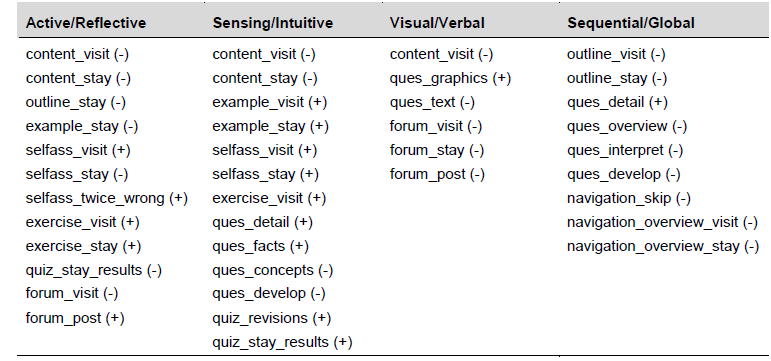
****

Fig 3. Questionnaire to measure the learner’s learning style [24]

*CGPA* is the academic measure of how a learner has been performing over the span of his degree. If student is taking course within the same degree, his CGPA in previous semesters is considered or that of his prior degrees is considered. (Assuming that learners are coming from computer science back ground).

*Pre-Req GPA* is another degree of performance that a person is exhibiting in pre-requisite course of the current course. (Assuming that “Fundamentals of Programming” is the prerequisite course of “Object Oriented Programming” course).

*Age* of learner is also an important factor for deciding upon learner-ability to grasp the ideas and concepts. It has been used to classify the learners as slower ones or quick learners.

*Locale/Origin* of learners has also been revealed as a factor with an impact on learner’s academic performance and completion of degree. Each of the cities is labeled with a score reference to metropolitan, semi-urban and under-developed.

*Learner Attributes* acquired from different sources have been used to categorize the learner profiles with respect to their level of expertise and skills. These learner categorizations can be used for subsequent recommendation of learning contents to the learner. Learner profiles are maintained in “LearnerOntology” as ontology classes, properties and individuals through the Protégé tool (with requisite restrictions on concepts) as shown in Fig. 4. Moreover, few of the restrictions on ontology concepts through Description Logic (DL) constructs are illustrated in Fig. 5. 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 3.1.

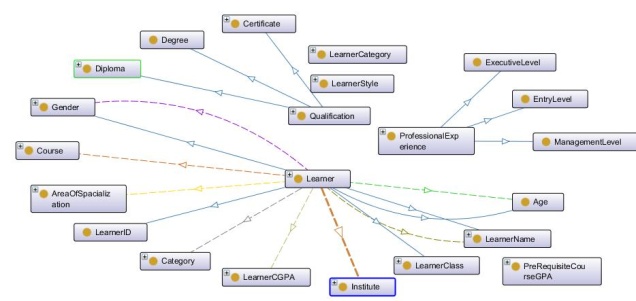
**

Fig. 4: A Snippet of LearnerOntology

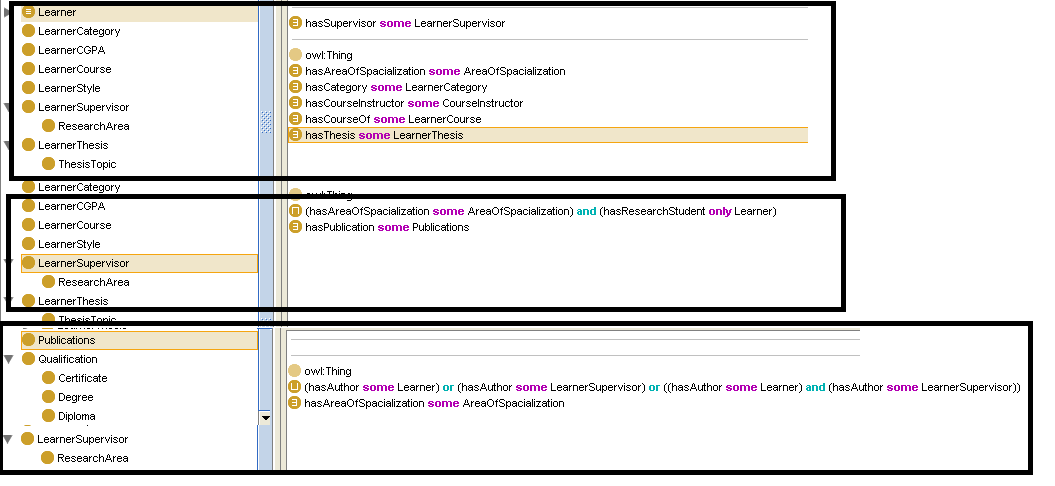


Fig. 5: DL Restrictions on concepts of LearnerOntology

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.

**3.2. Learner Categorizer - *LCHAIT***

In order to categorize the learner profiles, a novel technique has been proposed named *LCHAIT*. *LCHAIT (Learner Categorization with Hybrid of Artificial Intelligence Techniques)* is hybrid of two machine learning techniques named Case based Reasoning (*CBR)* [27] and Artificial Neural Networks (*ANN)* [28]*.* For new learners, profiles are retrieved in the same fashion as in the CBR retrieval phase using certain similarity metric [29]. Among the retrieved ones, cases that can be utilized are reused and the rest of them may be adapted for usage. The ANN model, used for categorization of learners through CBR’s adaption phase [30], is trained over the retrieved cases instead of trained on learner profiles contained in whole case base. Once the learner category is assigned to the learners into any of the four categories by the ANN, the profile is retained into the CBR’s memory for re-use.

Case based Reasoning (CBR) has been selected due to its inherent capacity of maintaining the prior knowledge into its repository and reusing the past knowledge maintained. In order to retrieve the most relevant cases the same similarity metrics were used i.e. Tvesky’s Ratio Model with same similarity threshold as previously. The business logic for component of CBR has been implemented using JAVA.

The variant of Neural Networks named MLP (Multi layer Perceptron) 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.

The ANN model has been trained over the data repository containing learner profiles as discussed in section 3.1. In order to train the ANN model, the dataset fed (i.e. the whole set of 600 profiles) was divided into three bins of training set, validation set and testing set. Training and validation phases were targeted for making adjustments to the ANN model in reference with its error rate and generalization. Subsequently, testing phase measured the model performance with respect to its accuracy. Moreover, it aids in deciding if the ANN model needs to be retained provided that the error rate exceeds that expected. The dataset of proposed model was divided into three sets with a division of 70% for training, 20% for validation and 10% for testing of the ANN model. 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

% Setup Division of Data for Training, Validation, Testing

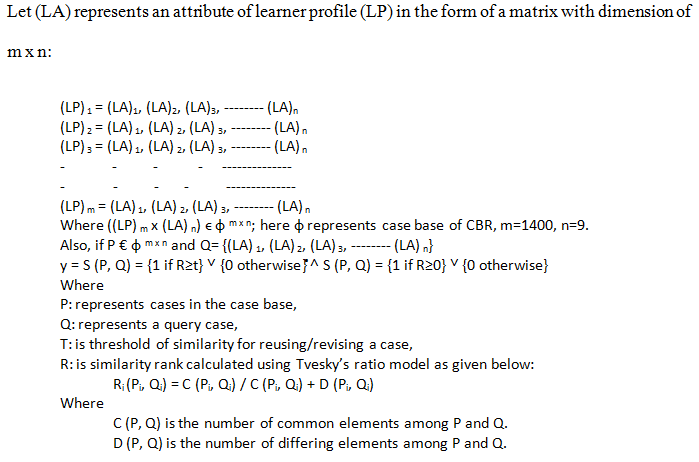
net.divideFcn = 'dividerand'; % Divide data randomly

net.divideParam.trainRatio = 70/100;

net.divideParam.valRatio = 20/100;

net.divideParam.testRatio = 10/100;

A mathematical representation of proposed approach has been given in the following to formally describe the process of learner categorization:



In our case, since the output is the discrete random variable with discrete time distribution; and if current output (i.e. selected profile); we can predict learner category ‘y+1’ which is function of current output ‘y(n)’ and profile values ‘ф(n)’ in the case base:

y (n+1) = f (ф(n), y(n))

For ANN model,

inputs are (y1,y2,y3-------yn) € {(LP)1 , (LP)2, - - - - - -,  (LP)m}

where:

(y1,y2,y3-------yn) ↔ (LP)

weights (w1, w2, - - - - wn) € W with one output

y(n+1) = ∑i=1 to n satisfying F: yk such that F(yi,wi) = f(yi, wi) ϵ ф

y(n+1) is the learner category predicted.

A linear model of input, state and output can be represented by:

dx/dt = Ax + By

y = Cx + Dy  ------------------- eq (1)

where y ϵ фmaximum with input x ϵ фmaximum  (xn) is state and y ϵ фmaximum **,** (xp) is the output.

Consequently,

A ϵ фn x n B ϵ фn x m  C ϵ ф P x n  D ϵ ф P x m

D is called feed through term. The polynomial form of eq(1) can be given as:

Ф(d/dt). Y = M (d/dt) x with y = cal (x,y)

* 1. **Learning Content Model**

In any learning management system, learner and learning content are the key factors contributing towards success of any system. Apart from maintaining the learner profile (as discussed in section 3.2), learning content modeling needs to be maintained as well. Learning content model is the representation of the properties of a course and more specifically one in object oriented programming languages along with contents.

Learning contents, represented by topics and subtopics each corresponding to a concept in ontology as shown in Fig 8, 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. Based upon performance of learner in certain topics and weeks, learning contents are re-sequenced or category of learner is updated. After half of the semester, learner’s performance is evaluated subject category in order to re-assign the learner category. The key parameters of adaptivity are the target-model scores for each of the category, course week, assessment score (weighted scores obtained in quizzes, assignments, exercises and class tasks) and number of times learner’s content level has been upgraded and downgraded while offering the contents.

Each of the topics is connected to sub-topics, pre-requisite topics and topics following the current ones (i.e. Post topics). Every topic is connected to a Learning Object (LO) that refers to repository containing learning contents in the form of text files, images or videos. Every learning object is annotated in such a way that it can identify and represent a topic in entirety as modeled in Fig 6 and 7 respectively.

LO111 means, it is a LO which 1, offered in week 1 with difficulty level of 1.

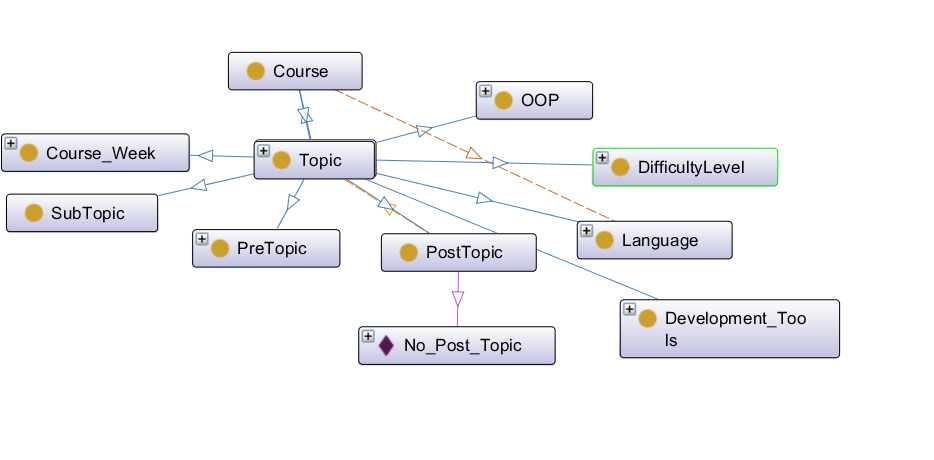


Fig 6: Domain representation of CourseOntology

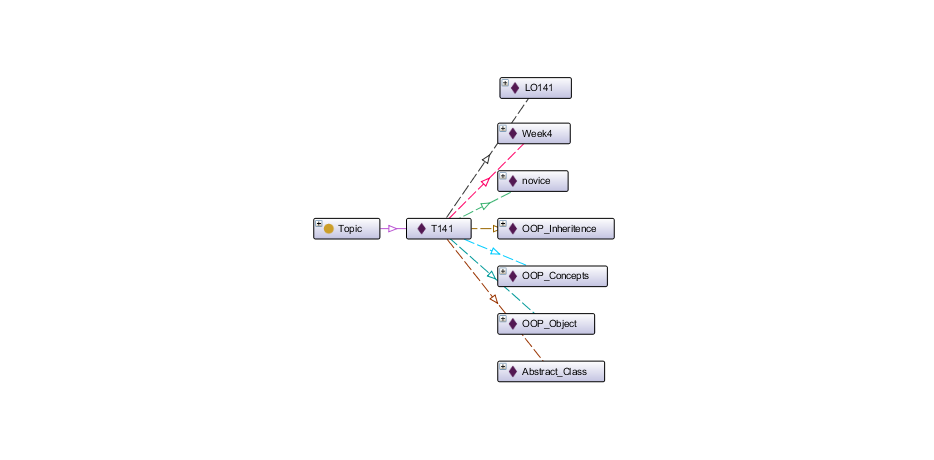


Fig 7: Topic with respective Instances and LO representation

Table 2: Representation of Typical Topics and LO as Ontology Instances

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Topic** | **Topic Name** | **SubTopic** | **Language** | **Week** | **Learning Object** |
| T111 | OOP | Objects | 1 | 1 | LO111 |
| Inheritance | 1 | 1 | LO112 |
| Abstraction | 1 | 1 | LO113 |
| Encapsulation | 1 | 1 | LO114 |
| T281 | Pointers | Declaring Pointers | 2 | 8 | LO291 |
| Ref/DeRef Pointers | 2 | 8 | LO292 |
| Array Pointers | 2 | 8 | LO293 |
| Function Pointers | 2 | 8 | LO294 |
| T3A1 | Threads | Threading | 3 | 10 | LO3A1 |
| Events | 3 | 10 | LO3A2 |
| Sync. Threads | 3 | 10 | LO3A3 |
| Semaphores | 3 | 10 | LO3A4 |

**3.4 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 or an expert level learner may be offered expert level contents (i.e. incorporating the feature of personalization).

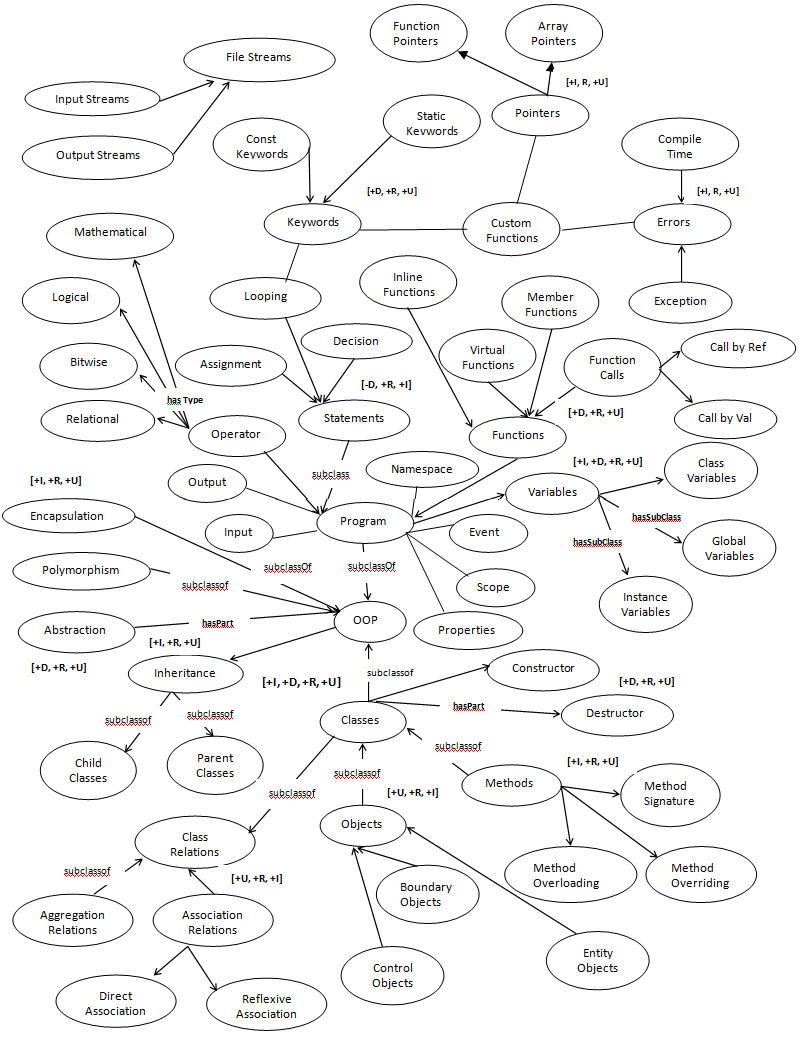


Fig 8: CourseOntology

Once learning activity is performed, learner is undergone the phase of assessment on bi-weekly basis. These assessments comprise of an average score learner has attained in assignments, quizzes and exercises. Moreover, learner’s performance is tracked over the weeks. If learner keeps on performing as desired (i.e. meets the target model criteria) for two weeks, contents that are offered to him are upgraded with a next level category of contents i.e. learner previously offered with novice level contents is presented with easy level contents and so forth. Likewise, if performance of a learner is not up to the mark (i.e. does not meet the target model criteria), he is down-graded to easier level of contents i.e. a learner offered with expert level contents is presented with proficient level contents and so forth. Furthermore, if learner’s performance does not perform up to the target model, category of the learner is downgraded or upgraded for presenting the respective learning content. This maneuvering of contents by sequencing/re-sequencing contents is driven by the “relative grading” that shuffles learners (in different categories) to ensure aspect of adaptivity.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Below Avg** | **Average** | **Above Avg** | **High** |
| **% Clusters** | Novice | Easy | Proficient | Expert |

Pseudo code for rule based semantic content recommendation (KASER) is given in the following:

Public Topic recommendLearningTopic (learnerID, learnerCat, cWeek) {

cAssessScore = getcurrentAssessmentScore(learnerID);

tModelScore = getTargetModel(learnerID,cWeek);

If (cWeek==1 or cWeek==2 && learnerCat==‘expert’) then

contentType = proficient;

If (cWeek>2 && cWeek<5) {

if (learnerCat==‘novice’ && currentAssessScore >= tModelScore) then

contentTopic=postTopic;

else if (learnerCat==‘novice’ && currentAssessScore < tModelScore) then

contentTopic= preTopic;

contentType = novice;

else if (learnerCat==‘easy’ && currentAssessScore < tModelScore) then

contentTopic= preTopic;

contentType = novice;

contentType = easy;

}

assignlearningContent based On:

[week and learnerCat and assessmentScore and targetModelScore]

Topic=pre-topic or Topic =post-topic

}

Public int getcurrentAssessmentScore(learnerID)

{

assessmentScore = [(assignmentScore + quizScore + exerciseScore)/ 3]

return %assessmentScore;

}

Public int getTargetModel(learnerID, cWeek){

category = getCategoryofLearner(learnerID);

Relative grading to decide the target.

Tgt=(Score - Mean)/SD

If (category) then

targetScore

return %targetScore

}

# 4. Results and Evaluation

# The evaluation of proposed system is performed from two perspectives i.e. Evaluation of Ontological models through standard ontology evaluation techniques and experimental evaluation of proposed techniques. The experimental evaluation of system has three aspects (a) Accuracy of learner categorization (b) Precision of KASER in content recommendation and (c) Impact of Proposed system on learner’s performance over a temporal scale.

**4.1. Ontology Evaluation**

There are two aspects of evaluating the health of ontologies; First, the ontologies which have been used in the learner categorization as well as in the content model and second is to compare proposed ontology model with contemporary ones [37, 38].

**4.1.2. Consistency of Ontological Models**

Ontology consistency is evaluated using Pellet [32] reasoner and built in ontology tester of Protégé (mainly checks the disjoint, inverse and domain/range). This reasoner ensured that ontology is consistent and can be utilized for semantic E-learning operations without creating any inconsistency among concepts/classes and their hierarchical relations.

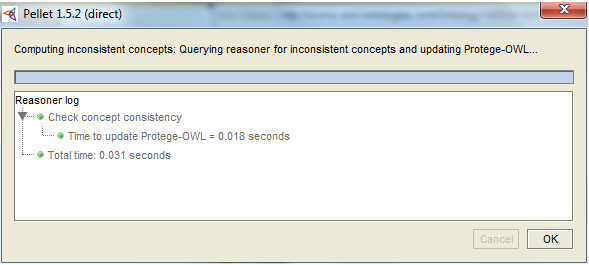


Fig 9. Consistency Check of Ontology Model

**4.1.3. Completeness (Domain Coverage) of Ontological Model**

The completeness metric of the ontological model determines the comprehensiveness of the domain coverage for E-Learning systems. Proposed ontology models cover almost all the concepts related to characterization of learner, domain of a typical university course, assessment modules and course of OOP. Moreover, ontology models comprise of necessary axioms ensuring redundancy-free terms and irrelevant axioms.

In order to verify the ontology coverage for fully modeling the domain of E-learning, SCORM [31] was used as a point of reference besides the competence of domain expert. SCORM enlists all aspects i.e. course content development, exams (assessment), quizzes, exercises, exams etc.

**4.1.4. Expandability and Reusability of Ontological Model**

The ontology model specified for the course i.e. CourseOntology, is expandable. This ontology can be expanded with new concepts to further elaborate the contents of course. Similarly, LearnerOntology may be updated with any new field to be added to build the profile of learner. Moreover, semantics-aware contents can be offered to the learners at all stages of learning cycle.

**4.1.4. Quality of Ontological Model**

Quality of Ontology evaluation refers to fault diagnosis and fixation of revealed gaps (or pitfalls). It is cohesive part of tasks involving ontologies for technically validating the ontology quality. In order to evaluate the overall quality of ontologies (in terms of faults) , two of the most prevalent techniques have been selected i.e. OOPS (Ontology Pitfall Scanner) [25] and Ontoclean [26].

A technique named OOPS (Ontology Pitfall Scanner) has been employed that diagnoses and categorizes the faults reference to their criticality fully aligned with quality standards. These categories are normal, minor, important and critical. OOPS findings for LearnerOntology are given in table 3 with no critical pitfall.

Table 3: OOPS based evaluation of LearnerOntology

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Pitfalls Category | Ontology Pitfall | Importance Level | | | |
| Normal | Minor | Important | Critical |
| Structural Dimension | Modeling Decisions- P[24] |  |  | √-1 |  |
| Wrong Inference | √ |  |  |  |
| No Inference P[11, 13] |  | √-4 | √-3 |  |
| Ontology language- P[34, 35, 38] | √ |  |  |  |
| Functional Dimension | Real World Modeling – P[4, 10] |  | √-3 | √-1 |  |
| Requirements Completeness |  |  |  |  |
| Application context- P[36, 37, 39] | √ |  |  |  |
| Profiling Dimension | Ontology Clarity – P [8,22] |  | √-2 |  |  |
| Ontology Understanding- P[8,11, 13] |  | √-3 | √-2 |  |
| Ontology Metadata – P[38, 41] | √ |  |  |  |
| Consistency | P[5] |  |  | √-1 |  |
| Completeness | P[4,10,11,13] |  | √-2 | √-3 |  |
| Conciseness | P[2,3] | √ |  |  |  |

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. The Figs 10 assert the fact of OntoClean which offers logic based argument in validating the ontological and taxonomical relationships through meta-properties of Rigid, Identity and Unity.

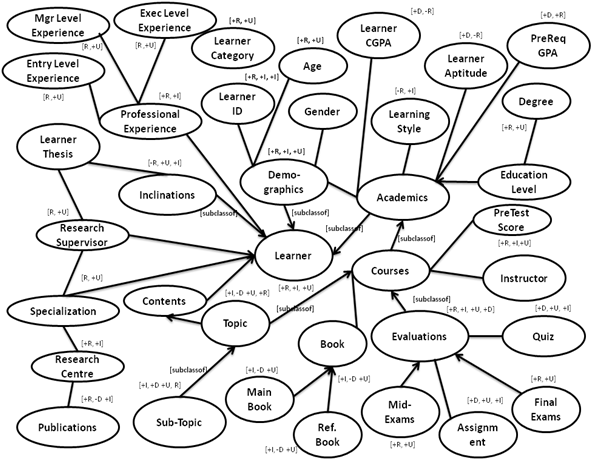


Fig 10**:** OntoClean based Evaluation of Learner Ontology

**4.2 Comparative Analysis of Ontologies (Domain Coverage)**

As discussed in section 3.4, a comparison of proposed ontology models (in KASER) has been made with ontological contents [32, 33] for comparison of ontology with respect to the coverage of domain (with prevalent techniques named SRS and PAeLS respectively).

In prevalent techniques specifically [32, 33], ontologies are focused towards personalization of presentational aspects than actual contents to be presented to the learner. Secondly, domain ontology is not very comprehensive to cover all aspects of a course in semester of a class. Lastly, the contents presented in ontologies are very specific and trivial with very few relations among concepts in terms of object properties and data properties. Our ontology, on the other hand, is comprehensive enough to cover all the requisite concepts for any programming language course in certain semester. Also, the aspects in content model contained in the form of ontology fully describe the object oriented programming languages at finer level of granularity (C++, C# and Java are modeled in ontology). Here a comparison is presented with respect to the coverage of domain i.e. number of concepts, number of object properties and the data properties as shown in Fig 11. The comparison in terms of coverage of concepts asserts the superiority of proposed technique for having better concepts coverage, properties and profiling of the learners.

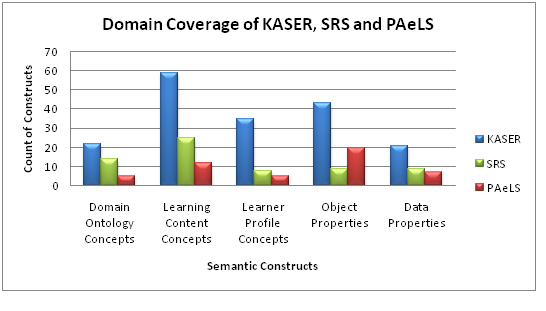


Fig 11. Comparative Analysis of Domain Coverage

**4.3 Evaluation of Learner Categorizer (*LCHAIT*)**

In order to measure the performance of *LCHAIT* and its comparison with existing knowledge base techniques i.e. Fuzzy Logic (FL) [34], Case based Reasoning (CBR) [35] and Neural Networks (NN) [36], a data set comprising of the profiles of 1000 students was used. In order to build the profiles, data of students was acquired from different universities (GCU Lahore1, UAAR Rawalpindi2, and Iqra University3), with different institutes (Computer Sciences and Management Sciences, Engineering Dept), and different courses (Fundamentals of Programming, Organizational Behavior, Data structures, Marketing and Business Law). This data diversity is purposefully introduced to comprehensively cover variety of cases in our data model.

Comprehensiveness of data models ensures an effective training approach of machine learning models independent of any biases (i.e. lack of coverage of cases or overfitting).

There were eight sets of new learners’ Profiles (each set having profiles of 50 learners) given randomly to retrieve similar profiles through CBR’s similarity metric as explained in section 3.4.

Out of 1000 learner profiles, 700 were used as training set of ANN model and 300 were used as validation set of ANN model in a random fashion (n-fold cross validation). Whereas, eight sets of new learner’s profiles were used for testing of ANN model. Effectiveness of ANN model was measured through its accuracy in assigning category to the learners presented as validation set.

*LCHAIT* has been evaluated through variation of training set and validation set with same testing sets as described in above scenarios.

1 <http://www.gcu.edu.pk/>

2 <http://www.uaar.edu.pk/>

3 <http://iqra.edu.pk/isl/>

Learner profiles retrieved through CBR’s retrieval phase are used for training of ANN model. Profile retrieval is carried out without provision of category, whereas retrieved cases contain the learners’ category. All eight sets (each having 50 profiles), serving as validation set, retrieve profiles (with 60% similarity) which are used to train the ANN model for accurately predicting category of learner.

A consolidated view of predictive accuracy exhibited by all four methods is shown in Fig 12, where every value is percent depiction for performance models. In other words, it is the picture of classification accuracy to rightly predict learner’s category without taking into account aspects of precision and recall.

Learner AccuracyPlot.eps

Fig 12: Comparative Analysis: Accuracy

Fuzzy Logic has shown an average accuracy of 29.67% in retrieving right cases. CBR has accuracy of 47.35% followed by accuracy of ANN i.e. 57.52%. *LCHAIT* has shown better accuracy than other techniques i.e. 70.84%. The reasons for superior performance of *LCHAIT* over Fuzzy Logic, Case based Reasoning and Neural Networks are discussed below.

Fuzzy logic, driven through knowledge in *Rule Inference Engine*, seems not adaptive to comprehend different scenarios with different parameter values due to static rule base. For example, profile attribute of *PreTest* has higher impact in classifying the learner to certain class but rules in the rule base cannot comprehend these relationships. Such variation in profile attributes are not handled adaptively by Rule-base of fuzzy logic.

The performance of Case based Reasoning is greatly dependent upon selection of right contents during retrieval phase through similarity metrics. Here, relevant profiles are selected based on static rank that inherently would ignore cases even if they are to be selected with a least margin without taking into account any exceptions.

Neural Networks exhibits better performance than FL and CBR due to its dynamic and adaptive nature. Besides input and output layers, there were two middle layers, each containing 15 neurons in order to get trained, validate and test the learner profiles for right categorization of new learners.

On the other hand, the hybrid approach named *LCHAIT* shows better performance than rest of the contemporary techniques in terms of performance parameters so far. *LCHAIT*, contrary to ANN, uses only one hidden layer with 8 neurons on middle layer (computationally less expensive). Effectiveness of *LCHAIT* is signified by phase of preprocessing (both for attributes and data), by data used for training and more importantly by training of NN model on most relevant profiles retrieved through CBR’s profile retrieval phase.

**4.3 Evaluation of Personalized Content Recommender (KASER)**

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 (*Knowledge based Adaptive Semantic E-Learning Recommender*)* keeping 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. During initial weeks, recommendation of contents while ensuring the personalization was done by rule-based recommender. Here degree of correctness for contents offered was main focus before assessing the learner. 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 4.

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 [37].

Table 4: KASER’s Personalized Content Recommendation Validated by Domain Experts (DE)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **No of**  **Weeks** | **Recommendations by KASER** | | **Accuracy Validation by DE** | | **Kappa ’s**  **Coefficient** |
| **Topics** | **Sub-Topics** | **DE 1** | **DE 2** |
| 1-2 | 3 | 5 | 72% | 81% | 74% |
| 3-4 | 4 | 9 | 83% | 77% | 79% |
| 5-6 | 4 | 7 | 80% | 85% | 81% |
| 7-8 | 3 | 5 | 83% | 68% | 72% |
| 9-10 | 2 | 6 | 76% | 84% | 78% |

Fig. 13 provides a detailed insight to the weekly progress that learners made while following the learner activity through conventional system, with PAeLs [38] and semantic recommender (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 2016 for course of "Object Oriented Programming".

It may be observed that learner’s performance is substantially improved when learning process is carried out by following the proposed approach (*KASER*). The results exhibited by *PAeLS* [38] appeared to be better than conventional approach; however, a visible performance difference can be seen compared to *KASER.* Few of the distinctive reasons for better performance of proposed approach are comprehensive number of parameters used for learner’s profiling, effective learner categorization through dynamic machine learning techniques, modeling of learning content at various levels of difficulty, personalized and effective content recommendation of contents based on learner ability and assessment based adaptivity of learning contents.

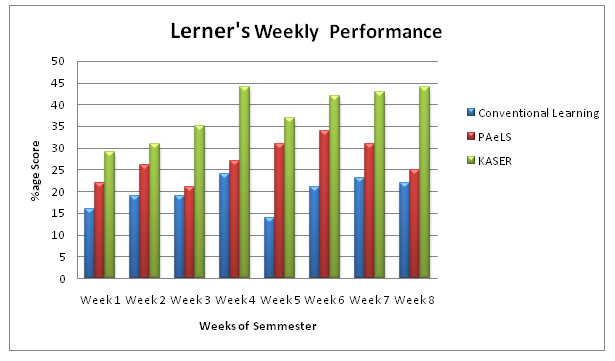


Fig 13: Learner’s Performance: Weekly Impact of KASER, PAeLs and Conventional Learning

**4.4 Impact of LCHAIT and KASER on Learner’s Productivity**

The evaluation process of proposed approach did not span over a week or a term rather learners from three batches in three years were categorized, underwent learning-assessment cycles and their performance was recorded. An aggregate of performance metrics for implementing the envisaged idea with personalization and adaptivity of learners was recorded. Fig. 14 provides a detailed insight to the year-wise performance of learners in mid-term and final-term exams. The average of these scores was taken in for 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 (i.e. conventional approach as shown in Fig 13).

In year 2016, with new group of students, learner categorization based content recommendation was made and subsequently results were recorded. The content recommendation was done with feature of personalization i.e. learning contents with respective level of difficulty for the relevant learners. Lastly, the results with features of assessment based content adaptivity were recorded in year 2017 on same patterns as in previous two years.

A consolidated view of the learner’s performances in mid-terms and final-terms of year 2015, 16 and 17 is illustrated in Fig 14.

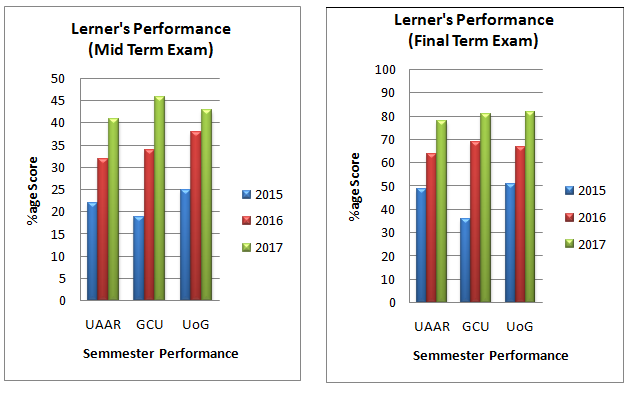


Fig 14: Learner’s Performance in Mid/Final Terms in Year 2015, 16, 17.

A thorough evaluation of proposed E-learning framework has been provided in preceding sections from two perspectives i.e. evaluation of ontological models and experimental evaluation of different dimensions of proposed framework. Ontologies have been evaluated based upon standard techniques such as OntoClean, OOPS, consistency and domain coverage as given in section 4.1. Ontologies used in modeling have been found complete, consistent and rich in terms of knowledge when compared to prevalent models (as given in section 4.2). The preliminary implementation of proposed framework shows an improvement in performance of learners as stated in section 4.2 to 4.4. The evaluation of techniques employed for learner profiling and categorization is explained in section 4.2 and 4.3. The results assert *LCHAIT* as a better technique than contemporary ones due to hybrid of neural networks and case based reasoning. The performance of knowledge based content recommender is evaluated in section 4.4 where accuracy of recommending the learning contents to learners is presented. Finally, an overall impact of learner’s performance is presented over different temporal scales i.e. learner’s weekly performance, performance in mid-term exams, improvements in final-terms exams and overall impact of proposed approaches on learning capacity of learners.

1. **Conclusion and Future Directions**

Keeping in view the technological advancements, potential increase in number of learners and their learning requirements; ontology based personalized and adaptive E-learning framework has been proposed. The rationale for proposed system has been provided in detail while justifying its alignment with prevalent technological trends (especially those of web3.0 technologies).

One of the most important goals of E-learning systems is its effectiveness for improvising learner’s productivity. So learner’s productivity has been ensured through features of personalization and adaptivity while suggesting the learning contents to the learner.

A comprehensive set of learner attributes was shortlisted in order to build the learner profiles. The selected profile attributes were further refined for having the ones with maximum impact on defining learner’s capabilities. Based on these profile attributes, each of the learners was assigned a category while keeping in view the cognitive skills of learners. A technique named LCHAIT was proposed that exhibited an effective degree of accuracy while categorizing the learners compared with rest of the techniques. The learning content annotated through ontologies and sequenced for learners of different category was recommended through a dynamic content recommender (KASER) with respect to learner’s category.

Overall, performance of learners was recorded in order to measure the impact of proposed approaches on semester- basis (weekly performance, mid-term and final-term exams) in years 2015, 2016 and 2017 respectively. The performance of learners appeared to improve with gradual incorporation of proposed approach compared with conventional approach and PAeLs. Such improvement is courtesy to comprehensive learner profiling, dynamic learner categorization, personalized content recommendation, adaptive learning cycle and reusable learning contents modeled in ontologies.

We look forward to use the fuzzy logic for clustering the learning contents and subsequently the learners based on their performance. Besides, experiments may be carried out for learner categorization using another variant of neural networks i.e. Radial Basis Function (RBF) [39] that works well with limited training/validation sets. Moreover, Genetic Algorithms (GA) would be used for learner categorization. Another dimension may be to experiment with fuzzy logic by making its rule base dynamic through GA as done by [40].

Moreover, we will develop a generic ontology backed system capable of developing contents dynamically from any domain.

**References**

1. Bates, T., Technology, e-learning and distance education, London: Routledge, 2005.
2. Ontology based Adaptive, Semantic E-Learning Framework (OASEF) Sohail Sarwar, Zia Ul Qayyum, M Safyan, Faisal Munir, Springer LNEE ICISA, 2016.
3. Bozkurt, A., Akgun-Ozbek, E. et al, Trends in Distance Education Research: A Content Analysis of Journals 2009-2013. International Review of Research in Open and Distributed Learning, 16(1),330-363. 2015.
4. Marian Cristian Mihăescu, “Classification of Learners Using Linear Regression”, Proceedings of the Federated Conference on Computer Science and Information Systems pp. 717–721 ISBN 978-83-60810-22-4. 2011.
5. Thakaa Z. Mohammad, Abeer M.Mahmoud, et al, Classification Model of English Course e-Learning System for Slow Learners “Recent Advances in Information Science”; International Journal of Computer Science (IIJCS) , ISBN: 978-960-474-304-9, 2014.
6. Van Seteres, M.A Ossevroot et al, Influence of student characteristics on use of adaptive e-learning material, Int journal of [Computers & Education](http://www.sciencedirect.com/science/journal/03601315) [Volume 58, Issue 3](http://www.sciencedirect.com/science/journal/03601315/58/3), Pages 942–952, April 2012.
7. Chusak Yathongchai et al, Leamer Classification Based on Learning Behavior and Performance, IEEE Conference on Open Systems (ICOS), Dec 2013.
8. Shute, V., & Towle, B., Adaptive e-learning. Educational Psychologist, 38(2), 105–114. 2010.
9. David P., A Design Requirements Framework for Distance Learning Environments, Journal of computers, Vol. 2(4), pp 99-113, 2007.
10. Darrell M. West, Mobile Learning: Transforming Education, Engaging Students and Improving Outcomes, International Journal of ICT, e-Management and e-Learning, Vol. 4, 2013
11. A case study on SCORM –based eLearning in computer aided drafting course with user satisfaction survey, WSEAS Transactions on Information Science and Applications, Vol 5(10), pp 1416-1427, 2008.
12. Luis E, Anido Rofio et al., “A recommender system for educational resources in specific learning content”, International conference on computer scirnce and Education, 2013
13. Salam, Fraihat, Qusai Shambour, “A Framework of semantic recommender system for e-learning”, Journal of Software, vol 10, 317-330, 2015.
14. Romero and Ventura S., “Educational Data Mining: A Survey from 1995 to 2005”, Expert Systems with Applications, 33(1), pp. 135-146, 2007.
15. Saleena B., Srivastava K., “Using concept similarity in cross ontology for adaptive e-learning”, Journal of King Saud University- Computer and Information Sciences, Vol 27(1), pp. 1-12, 2015
16. Parneet Kaur, “Classification and prediction based on DM algorithm for slow learners”. International Conference on Recent Trends in Computing 2015.
17. Paolo Bouquet, Andrea Molinari, Using Semantic Technologies in E-Learning Platforms: a Case Study, International Journal of Information and Education Technology, Vol. 6(5), 2016
18. Monika Rani, Kumar Vaibhav Srivastava, Om Prakash Vyas, An Ontological Learning Management System, Computer Applications in Engineering Education, Vol. 24(5), pp. 706-722, 2016
19. Sarma Cakula, M Sedleniece, “Development of Personalized e-learning model”, ICTE in Regional Development, Vol 26(4), pp. 113-120, 2013.
20. Maryam Yarandi, Hossein Jahankhani, Abdel-Rahman H. Tawil, “A personalized adaptive e-learning approach based on semantic web technology”; Webology, Volume 10, Number 2, December, 2013.
21. https://www.scribd.com/document/165064823/GAT-Subject-Syllabus, GAT Test Bank Accessed on Oct 2016.
22. Robert Lafore, Object-Oriented Programming in C++, Fourth Edition,  Sams Publishing, Published December 29th 2001.
23. <https://www.tutorialspoint.com/cplusplus/cpp_online_quiz.htm>, OOP Quizes, Accessed on Sep 2015
24. Silvia Rita Viola , Sabine Graf , Kinshuk, Tommaso Leo,” Analysis of Felder-Silverman Index of Learning Styles by a Data-driven Statistical Approach”, IEEE International Symposium on Multimedia (ISM'06), pp. 7695-7704, 2006.
25. María Poveda-Villalón et al., “Validating Ontologies with OOPS!”, The 18th International Conference on Knowledge Engineering and Knowledge Management (EKAW), pp. 267-281, 2012
26. N.Guarino and C. A. Welty,”An Overview of OntoClean”, The Handbook on Ontologies Berlin:Springer-Verlag, 2004
27. D. Arditi, B. Tokdemir, Comparison of Case-Based Reasoning and Artificial INeural Networks, Journal of Computing in Civil Engineering, Vol 13(3), pp. 578-583, 1999.
28. L. Shen, R. Shen, Ontology based Content Recommendation, International Journal of Continued Engineering and Education, Vol 15(1), pp. 13-26, 2005.
29. A. Aamodt, E. Plaza, Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches” AI Communications, Vol. 7(1), pp. 39-59, 1994.
30. M. Safyan, Z. Qayyum, S. Sarwar, Context-Aware Personalized Activity Modeling in Concurrent Environment, Internet of Things (IoT), 2017.
31. B. Liem, W. Beek, J. Gracia, E. Lozano, [DynaLearn–An Intelligent Learning Environment for Learning Conceptual Knowledge](http://www.aaai.org/ojs/index.php/aimagazine/article/view/2489), AI Magazine Vol. 34 (4), pp. 46-65, 2013.
32. M. Yarandi, A Personalized Adaptive E-Learning Approach Based On Semantic Web Technology, Journal of Webology, Vol. 10(2), pp. 751-766, 2013.
33. B. Liem, W. Beek, J. Gracia, E. Lozano, [DynaLearn–An Intelligent Learning Environment for Learning Conceptual Knowledge](http://www.aaai.org/ojs/index.php/aimagazine/article/view/2489), AI Magazine Vol. 34 (4), pp. 46-65, 2013.
34. S. Tambe, G. Kadam, An Efficient framework for E-Learning Recommendation system using fuzzy Logic and Ontology”, International Research Journal of Engineering and Technology (IRJET), Vol. 3(6), pp. 2062-2067, 2016.
35. S. Sarwar, Raul García-Castro, Zia Ul Qayyum, Muhammad Safyan, Ontology-based Learner Categorization through Case Based Reasoning and Fuzzy Logic, International Conference on E-Learning (IADIS), pp. 159-163, 2017.
36. S. Sarwar, Zia-ul-Qayyum, Raul Garcia Castro, M Safyan, “Ontology based E-learning Systems: A Step towards context aware content recommendation”, International Journal of Information and Educational Technology, Vol. 8(10), pp. 10-19, Oxford England, 2018.
37. S. Vanbelle, A New Interpretation of the Weighted Kappa Coefficients, Psychometrica, Vol. 81(2), pp. 399–410, 2016.
38. H. Jahankhani, R. Tawil, Adaptive E-learning Approach based on Semantic Web Technology, International Journal of Webology, Vol. 10(2), 2015.
39. P. Strumiłło, W. Kamiński, Radial Basis Function Neural Networks: Theory and Applications, Neural Networks and Soft Computing. Advances in Soft Computing, Vol 19(5), pp. 107-119, 2013
40. S. Tambe, G. Kadam, An Efficient framework for E-Learning Recommendation system using fuzzy Logic and Ontology”, International Research Journal of Engineering and Technology (IRJET), Vol. 3(6), pp. 2062-2067, 2016.