Knowledge Enriched Learning by Converging Knowledge Object & Learning Object

Sai Sabitha1, Deepti Mehrotra2, and Abhay Bansal2
1Research Scholar, Amity University, Uttar Pradesh, India
2Amity University, Uttar Pradesh, India
saisabitha@gmail.com
mehdeepti@gmail.com
abhaybansal@hotmail.com

Abstract: The most important dimension of learning is the content, and a Learning Management System (LMS) suffices this to a certain extent. The present day LMS are designed to primarily address issues like ease of use, search, content and performance. Many surveys had been conducted to identify the essential features required for the improvement of LMS, which includes flexibility and a user centric approach. These features can suffice the need of all learners, when they have different learning requirements. For a true learning, knowledge should also be delivered along with the domain information. There is a need to design an architecture for user centric Knowledge Driven Learning Management System (KDLM). Thus, for holistic learning, knowledge enriched teaching skills are required, which can enhance and increase the thinking skills of the learner to a higher level. The current LMS needs an improvement in the direction of knowledge discovery, exploration so that knowledge enriched learning can be provided to the learner. It can be based on knowledge engineering principles like ontology, semantic relationship between objects, cognitive approach and data mining techniques. In this paper, we are proposing an idea of an enhanced Learning Object (LO) called Knowledge Driven Learning Object (KDLO), which can be delivered to the user for better learning. We had used a data mining approach, classification to harness, exploit and classify these objects according to their metadata, thereby strengthening the content of objects delivered through the LMS.

Keywords: LMS; KMS; Learning Object; Knowledge Object; Classification; Decision Tree; Knowledge Driven Learning Objects; Knowledge Driven Learning Management System; e-Learning.

1. Introduction
At the beginning of 2001, the major issues that e-Learning communities were facing are resource sharing, repurposing, and inter operation between different e-Learning systems. These issues led to the growth of two kinds of e-Learning systems, LMS and Learning Content Management System (LCMS).

LMS is a high-level strategic solution that handles all aspects of the Learning Process and a portal for collaborative learning (Greenberg 2002). LCMS is primarily used for management of Learning Contents. It has elements like an authoring and assembly tool for creating Learning Objects, a repository for storing them and a dynamic engine for delivering Learning Content to the learners. The development of LCMS can be considered as a great progress for e-Learning environment, where technology of creating and maintaining Learning Objects was standardized. Till, here the Learning Object is considered as a core concept of Learning Content and thus LMS were mainly informatics centric rather than knowledge centric.

According to Experts & Training Industry (2010), LMS also plays a pivotal role in all business domains. The current LMS and LCMS were developed basically for the delivery of study materials, but today there is a need to develop an environment that helps instructors and students in value added teaching/learning, with a focus on what students need to know. The LMS should also support a situation where the learner has some preferred Learning Style (Graf & Liu 2008). Definitely, there is a need of wider metadata so that learners can select the content as per his choice. These issues were resolved to an extent by developing an LMS based on web 2.0, which helps in collaborative learning and creation of dynamic digital repositories. These Learning Systems have built-in collaboration tools (Blogs, wiki), which provide a space for learners to interact, share and learn through mutual collaboration. Many of them also provide tools for interaction with instructors, mentors and peers via a discussion forum, chat and virtual communities.

Today the K-economy is the buzzword frequently used by the people from all the domain of business. This has shifted from product based economy to a knowledge-based economy resulting in an increased demand for the workers who are capable of higher-order thinking and reasoning in solving intricate and authenticate problems.
Sai Sabitha, Deepti Mehrotra, and Abhay Bansal

in the workplace. All leading business organizations are now working on the principle of knowledge management where knowledge of its people is considered the most valuable resource of the organization. Knowledge Management (KM) is essentially about facilitating the processes by which knowledge is created, shared and used in organizations (Servin & De Brun 2005).

If we summarize the current need, it can be easily identified that there is a need to have a learner-centric environment in which the learner is not just a passive recipient of information but an active participant in knowledge acquisition. The major concern that remains in an e-Learning environment is how, to improve the process of knowledge acquisition. The role of an instructor or teacher needs to be modified such that knowledge together with domain information can be delivered to Learners.

Currently the science of learning is not only focusing on applications which transfers repeatable content, but applications which deliver the content in a new context and LMS in the future would have the ability to offer a richer content and a new learning experience for the user. This may help with retention of information and application of knowledge into work, leading to the development of new skills and deeper approach towards learning. Further, it increases the focus on understanding the content delivered rather than simply relying on memorization of the course material.

One way to generate knowledge is to design Knowledge Management System (KMS). Through these systems, small and relevant information about any domain is created. They are referred as knowledge nuggets in this paper. However, these knowledge nuggets are usually unstructured and hence cannot be a part of LMS directly. We need to convert these knowledge nuggets to structured objects called Knowledge Objects (KO) using an expert system. These KOs can also be used in a learning environment.

A classification technique, is adopted in this paper, which allows KOs to get associated with LOs. Thus, KOs are classified with different LOs and the resultants LOs can be considered as Knowledge Driven Learning Objects. Through these KDLOs, the learners receive requisite Knowledge. We had used decision tree algorithm for classifying the KOs with respect to LO. The basic idea of this research is to deliver a more valuable learning asset to the student and offer a right learning experience to each student in an e-Learning environment.

2. Background

2.1 Learning objects

The Learning Objects (LO) are now considered as a fundamental element of a new conceptual model for content creation and distribution. They can be defined as a reusable chunk of data that can be used as a modular building block for e-Learning content. The various structures of LOs were discussed and proposed (Wagner 2002; Wiley 2002; SCORM 2005; Heins & Himes 2006; Metros 2003). Reuse and repurpose of the LOs enables an LO to be used in different ways and in different programmes (Watson 2010; Collis & Strijker 2004; Gunn et al 2005; Polsani 2006). These objects stored in the Learning Object Repository (LOR) are fetched using metadata. There are many widely available metadata standards. (IMS 2006; Dublin core 2012; IEEE LTSC 2002; SCORM 2005) Among these Instructional Management System (IMS) and Sharable Courseware Object Reference Model (SCORM) handles both metadata specification and content structure modelling.

2.2 Learning management systems

LMS has been widely used in higher learning institutions as a mechanism to aid teaching and learning process. LMS provides an infrastructure and a platform through which learning content is delivered and managed (Balki 2010). The software tools of LMS have a variety of functions related to online and offline training, administration and performance management. It provides interaction between learner, instructors and content. The core modules of any LMS are student registration, course enrolment, course delivery, student performance tracking and student assessment. The LMS uses standardization in its construction and many organizations had developed regulations or recommendations about it.

2.3 Knowledge object

To extract knowledge in any organization tools and techniques of KMS is used. Knowledge Management (KM) is facilitating the processes by which knowledge is created (Tacit knowledge), shared and reused in organizations, so that an organization can obtain the greatest value from the knowledge available to it. The
Tacit Knowledge in a Knowledge Conversion Process can be considered as the content for Knowledge Object. Merrill (1999;2000) defines a Knowledge Object as:

“A record of information that serves as a building block for a Knowledge Management System. It has a content, method of organizing a knowledge base (metadata), rules to identify and categorize Knowledge Components.”

According to Horton (2001):

“A Knowledge Object is a chunk of electronic content that can be accessed individually and that completely accomplishes a single goal.”

Knowledge Object contains elements, like goal, content and support metadata (summary, introduction, keywords, related knowledge object, security information). A KO also represents information that has been semantically conceptualised (Ruffner & Deibler 2008). Finding ability and reusability are two important features to be taken into consideration while creating a Knowledge Object. This finding ability can be achieved by a metadata and reusability by using a Knowledge Object in various learning contexts and with the advancement of IT technologies, it is possible to dynamically create and distribute the knowledge and make it as a part of the learning mechanism. The various techniques used in extraction of knowledge in an organization are Communities of Practice, Sharing Best Practices, Knowledge harvesting, Peer assists, Social network analysis, After Action Reviews (AAR), Storytelling (Servin & De Brun 2005).

An educational organization involves various learning processes and this information can be collected from course plan, lesson plan, assignments, quizzes, tutorials, seminar etc. This data is usually uploaded on the portal of the college or organization. Some explicit knowledge can be generated through the quantitative feedback from students/learner through feedback management system and result analysis. Some of the tacit knowledge extracted within an educational organisation is achieved through the collection and storage of research projects in project management system and the creation of new knowledge and extraction through innovation management system. These are various forms of knowledge nuggets within the organisation.

Today’s e-Learning environment consists of several components of web 2.0 applications such as wiki, web logs, social book and Really Simple Syndication (RSS) feeds. The aim of the wiki was to develop an easy to use KMS enabling effective and efficient online collaboration. They are also apt for preserving and organising knowledge. Web logs contain rather simple units of information, permitting a more agile management of information. Bookmarks relating to any kind of web resource are stored in databases and social book marking generates a comprehensive and thoroughly indexed collection of scholarly resources. (Blees & Rittberger 2009). These also can be chosen as sources of knowledge nuggets after evaluation from an expert panel.

3. Proposed problem

3.1 Conversion of knowledge nuggets to knowledge object

The Knowledge Nuggets which are extracted from various systems need to be evaluated by an expert panel and get converted from an unstructured entity to a structured entity. This can be achieved by encoding it with a semantic relevant description of a particular object. Thus, adding a learning objective and requisite metadata these knowledge nuggets are converted to a KO which forms the building blocks of knowledge base, an integral part of the KMS. These conversions can be done through an expert system of KMS or by Ontology and are shown in Figure 1.

![Conversion of knowledge nuggets to knowledge object](image.jpg)

**Figure 1** Conversion of knowledge nuggets to knowledge object.
3.2 Proposed Knowledge Driven LMS

An LMS is a static system with defined tools for all learners. They deliver the content framed by a teacher or contributor. Today a more responsive and personalized experience of learning is needed by the user. So ideally the material that the learner receives must be blended with core knowledge of that subject saved as a knowledge object in the knowledge base of KMS, permutation of content objects which are stored as LOs in LMS. The idea proposed here is to develop an enhanced LMS, which has LOs and can be combined with extracted KOs from KMS to form as KDLO (Knowledge Driven Learning Object which can be delivered to the learners. These learning contents, make an effective use of various distributed knowledge, which enriches learning experiences of the learner or we can say a Knowledge-pull occurs in a knowledge driven LMS, where, along with the learning content, the knowledge is pulled and the need of a particular learner can be fulfilled to a great extent. Formation of a KDLO can be achieved by natural language processing, artificial intelligence, ontology and data mining. Classification, a data mining technique is used for the formation of KDLO.

3.3 Granularity

Granularity defined by IMS is, ‘the relative size of the resource’. According to IEEE, ‘It is an aggregate level or the functional size of the resource’. Granularity in the context of the LOs is often used to refer to the size of an object. Different scholars had discussed theoretical account of the aggregation level of a LO (Thompson & Yonekura 2005; Balatsoukas et al 2008; Wiley, Gibbons & Recker 2000). The term granularity refers to an object with the smallest level being a picture or a text and the largest being a set of courses. In a static e-Learning environment, a full course would be considered as grain size. However, in a dynamic LMS environment, we need a variety of objects so that more options can be delivered to the learner, and a semantic and cognitive approach can be used for identifying the most relevant object for the user. According to Hodgins (2002), a relevant object is one of the key strategies that determines how successful a LO is. A finer level of granularity ensures a greater potential to reuse of objects. Porter 2001, describes the notion of modules or modularity as “a key to provide flexibility for learners to use objects”. Designing learning as a modular concept in contrast to a course concept allows us to move from the ‘traditional course-building approach’ to that of ‘building-block concept’. In this paper, we had considered an environment where learners have a wider choice to select a LO for a particular subject and also there are a number of small KOs which can be attached to LO. The resultant KDLO also enhances the knowledge delivered with LO. However, for the sake of simplicity, relative size of a KDLO can be considered as one or two.

3.4 Proposed Model

![Figure 2: Formation of KDLO](image)

3.4.1 Model Description

The proposed model of KDLO is given in “Figure 2”. The model can be explained in the following steps.
1. Generation of LO and Metadata in LMS.
2. Extraction of Knowledge nuggets from the user through KMS and the conversion of nuggets into KO by adding a goal or an objective and Meta Data.
3. Convergence of LO and KO is done through Classification algorithm of data mining.
4. The decision tree classification technique is used for getting the KDLO as shown in Figure 2. Here, the Metadata of both LO and KO are considered for the classification algorithm.
5. For each LO we may have one or more associated KO which can be further considered as a part of Instructional unit.
6. Based on the user need the relevant KDLO can be delivered.

3.5 Data Mining Technique:-Classification

The aim of data mining process is to extract implicit knowledge from large volumes of data sets and transform it into an understandable structure for further use. These techniques are being used in various domains including educational domain. Data mining contains several algorithms and techniques for finding interesting patterns from large data sets. These are classified into two categories, namely, supervised learning (classification) and unsupervised learning (clustering) (Tan et al 2006). The classification algorithms of data mining are decision tree, rule based classifier, naïve Bayes classifier and K-Nearest Neighbour.

Classification is the task of learning a target function that maps each attribute set ‘x’ to one of the predefined class labels ‘y’. A Classification Model is useful for both descriptive and predictive modelling. A descriptive modelling can serve as an explanatory tool to distinguish between objects of different classes. In this approach, the metadata of the LMS has more attributes than that of KMS, thus the classification model can suitably be used to predict the class label of unknown records. Classification techniques are also most suitable for predicting or describing data with binary or nominal categories.

3.5.1 Need of Classification Technique –Decision Tree

We chose a Decision tree (Wu et al 2008) structure, which is self-explanatory as it can handle high dimensional data set (There are 63 attributes, which were identified by IEEE in the metadata standard of LMS). They are computationally inexpensive, and we can construct models quickly, even for a large data set (LMS and KMS are definitely a huge data set where the new LO and KO can be added as and when created or needed). As LMS and KMS may have redundant attributes, choice of decision trees is further justified as the presence of redundant attributes does not affect the accuracy of a tree. If an attribute is redundant with other, then one of the two attributes will not be used for splitting once the other attribute is chosen. Some other characteristics (Tan et al., 2006) which support our choice to select Decision Tree are:-

- Decision tree induction is a non–parametric approach for building model. It does not require any prior assumptions regarding the type of probability distribution satisfied by the class attribute or other attributes.
- Decision tree uses a heuristic based approach to guide their search in hypothesis space.
- Classifying a test record is fast with a worst-case complexity O (w) where w is the maximum depth of the tree.
- The decision tree is quite robust to noise when methods of over fitting are employed.
- As a result of classification, one or two KO can be classified to LO. It is represented as Knowledge Driven Learning Object. A delivery engine will deliver the appropriate KDLO to the user based on his or her requirements.

3.5.2 Algorithmic approach of Decision Tree (Han et al 2006)

Input: The training samples D, samples, represented by discrete-valued attributes; the set of candidate attributes, attribute-list.

Output: A decision tree.

1) Create a Node N

2) If <Tuples in data set D belongs to the same class>

then label the node with class C and return N as the leaf Node
3) If the attribute list is empty, then label the node with the majority of class C and return N as the leaf node.

4) Apply **Attribute selection Method (d, attribute)** to find the best splitting criteria.

5) Label node N with splitting attribute and branches for the outcomes as ‘J’
   (If splitting criteria are discrete multi way split)

6) New attribute list = attribute list – splitting criteria

7) For each outcome j of the splitting attribute, let D_j be the set of tuples satisfying outcome J.
   If D_j is empty
   {then attach a leaf node with majority class in D to node N}
   else if D_j be is not empty
   {then generate decision tree (D_j, new attribute list)
   attach a node return by generate decision tree to N}
   end for.

**Attribute selection Method (d, attribute ‘A’)**

- An attribute selection measure like info gain or Gini index is applied to the attribute.

- If ‘A’ is a discrete value (the outcome of the attribute A) for test node N then for having each known value of ‘A’ a branch is created.

- If ‘A’ is a continuous value (the outcome of the attribute ‘A’ has two possible values) for test node N then the value of ‘A’ has a split point and two outcomes (A<=splitpoint) and (A>=splitpoint).

- If ‘A’ is a discrete value and a binary tree to be produced, then returns a splitting subset of ‘A’ which has two outcomes

- Return Splitting Criteria

4. **Experimental Setup**

Dublin Core metadata contains 15 metadata attributes and IEEE Metadata standard has more than sixty attributes.

- Dublin Core uses metadata like title, subject, contributor, date created, type, etc. and these attributes are considered for LOs.

- KO can have metadata like title, author, date created, time, knowledge source, patent, knowledge type, knowledge objective.

- In the data cleaning step, set of five common attributes has been identified based on entropy measure and are considered for both KO and LO.

- They are an object id, title, author name, topic and sub topic.

- A data set of the 100 Learning Object Metadata is considered during the training phase and a set of 15 of KOs are considered for the test phase.

- Rapid Miner (RM) Tool is used for classification using decision tree.

Following steps were performed:-

Step 1: LOs and KOs metadata are loaded into rapid miner as shown in Figure 3.
Step 2: A Simple Validation tool as shown in Figure 4 randomly splits up the example set into two parts, ie, training set and a testing set. The purpose of the training set is to create models, whereas the testing set is used to estimate the accuracy of the created model. Here cross-validation operator of RM is used. Cross-validation is a standard statistical method to estimate the generalization error of a predictive model. In the training phase, the cross-validation is built on the current training set.

Step 3: A Decision tree operator of RapidMiner is used to classify nominal data on the training part as shown in Figure 5.

Step 4: Apply Model operator is used in the testing part. Models obtained after the training part usually contain information about the data on which they had been trained on. This information can be used for predicting the value of a possibly unknown label (Refer Figure 5).

Step 5: Validation allows us to estimate the accuracy of our model and Rapid Miner provides a tool, Performance Operator. (Refer Figure 5)
5. Result & Analysis

As there are more than two categories for the attribute used as the class label, C4.5 algorithm is used for the classification. A gain ratio is evaluated to decide the splitting criteria of the variable. Here in the experiment the class label taken is ‘object id’, which has a number of categories. The formation of KDLO proposed in Figure 2 is achieved using the data set and the output is shown in Figure 6. The KO1 and KO3 are classified together with LO9 for the data as follows:-

- Topic: Data Mining Theory.
- Sub Topic: Association.

5.1 Text View of DT

According to the “Text view” of Decision Tree shown in Figure 6, the KO1 & KO3 has been classified with LO9 based on Metadata attributes “topic” and “subtopic”.

For the given attributes, topic “dm theory” and subtopic “association” classifies a Learning Object (LO9) with KO1 and KO3. The output of KDLOs is further explained in “Figure 6a”.

Figure 5 Decision tree tool in Rapid miner

Figure 6 Decision Tree in text view
5.2 Graphical View of DT
Figure 7 shows the decision tree in a graphical form. The graph shows the attributes like title, topic, sub topic and author name form the nodes. The attribute “topic” has the highest entropy value and it is the root node. The object_id was chosen as the label, they are the leaf nodes.

Figure 7 Graphical view of Decision tree

5.3 Results of Performance Vector
In the testing phase, the accuracy of the decision tree is computed on the test data set as shown in Figure 8. The accuracy of a performance vector is 55%. (The arithmetic mean is ‘54.74%’ and ‘8.48%, is standard deviation).
Figure 8 Performance Vector.

6. Conclusion and Future Scope

By converging KO with LO, during the routine learning a learner gets a bigger picture of the topic to learn. It is one way to reinforce the value in learning, which can build knowledge elements into training programmes. Here we integrate the static learning resource (LO) and dynamic knowledge (KO) during learning, thereby making the learning operations efficient, effective, flexible, high accessible, relevant (Hodgins 2002; Ochoa & Duval 2008) and collaborative. In the above experiment, we have shown one or two KOs classified with a LO. The Knowledge Driven LMS will deliver a LO that can be enriched with the knowledge. According to the experiment, when a user queries for a learning material under course –‘data mining’, and module-‘dm-theory’, the LO with id- ‘LO9’ is delivered by LMS. Along with it, KO1 and KO3 are also given. Thus, KDLO can be obtained and thereby it can strengthen the contents of objects delivered through the LMS.

We had used decision tree algorithm for classification of LO and KO. The result shows, the KOs getting associated with one of the most relevant LOs. The other classification algorithms like naive Bayes classifiers, k-nearest neighbour and artificial neural network can be used. Other approaches can be done by using clustering and followed by classification. The data mining techniques like clustering algorithms (K-mean, Density Based Scan) can also be used. Text mining techniques which may identify the similarity index between different documents will definitely give a more relevant result. These techniques may further refine the search but at the same time increase the computational cost. However, the need of reusability of KO (Ruffner & Deibler 2008) is very important and can be achieved if we associate the Knowledge Object with more relevant LOs. Relevant and reusable objects can be achieved using Fuzzy Clustering, where an object can belong to more than one cluster. Associating relevance and ranking to the learning objects (Sabitha et al. 2012) can improve the satisfaction level of the end user using the LMS. However the delivery engine can be further modified using the semantic and cognitive approach that justifies the need of the user.

References


Wagner, E. (2002). Steps to creating a content strategy for your organization. Best of the, 103.


