International Journal of Future Generation Communication and Networking
Vol. 13, No. 4, (2020), pp. 5099 - 5108

A Study of Data Mining Procedures for Unification of Learning Object

Sami Abdul Qader Mahammed Al - Ademi
Research Scholar, Department of Computer Science, Dr. A.P.J. Abdul Kalam University, Indore, M.P.

Dr. Sanjay Singh Bhadoriya
Research Guide, Department of Computer Science, Dr. A.P.J. Abdul Kalam University, Indore, M.P.

Abstract

Data mining is one of the most important steps of the knowledge discovery in databases process and is considered as significant subfield in knowledge management. Research in data mining continues growing in business and in learning organization over coming decades. This review paper explores the applications of data mining techniques which have been developed to support knowledge management process. The main accomplishments of the research were the gathering of SOA functional requirements and the development of a prototype that provided an approach for the resolution of the interoperability gap that exists between LMSs and LORs. The article first briefly describes the definition of data mining and data mining functionality. Then the knowledge management rationale and major knowledge management tools integrated in knowledge management cycle are described.

Keywords: *Data Mining Procedures, Learning Object, Data mining, Databases Process.*

1. Introduction

A learning object is "a collection of content items, practice items, and assessment items that are combined based on a single learning objective". The term is credited to Wayne Hodgins, and dates from a working group in 1994 bearing the name. The concept encompassed by 'Learning Objects' is known by numerous other terms, including: content objects, chunks, educational objects, information objects, intelligent objects, knowledge bits, knowledge objects, learning components, media objects, reusable curriculum components, nuggets, reusable information objects, 2 reusable learning objects, testable reusable units of cognition, training components, and units of learning.

The Institute of Electrical and Electronics Engineers (IEEE) defines a learning object as "any entity, digital or non-digital, that may be used for learning, education or training". Chiappe defined Learning Objects as: "A digital self-contained and reusable entity, with a clear educational purpose, with at least three internal and editable components: content, learning activities and elements of context. The learning objects must have an external structure of information to facilitate their identification, storage and retrieval: the metadata." The following definitions focus on the relation between learning object and digital media. RLO-

ISSN: 2233-7857 IJFGCN

CETL, a British inter-university Learning Objects Center, defines "reusable learning objects" as "web-based interactive chunks of e-learning designed to explain a stand-alone learning objective".

Daniel Rehak and Robin Mason define it as "a digitized entity which can be used, reused or referenced during technology supported learning". Adapting a definition from the Wisconsin Online Resource Center, Robert J. Beck suggests that learning objects have the following key characteristics: Learning objects are a new way of thinking about learning content. Traditionally, content comes in a several hour chunk. Learning objects are much smaller units of learning, typically ranging from 2 minutes to 15 minutes.

Are self-contained – each learning object can be taken independently Are reusable – a single learning object may be used in multiple contexts for multiple purposes Can be aggregated – learning objects can be grouped into larger collections of content, including traditional course structures 3 Are tagged with metadata – every learning object has descriptive information allowing it to be easily found by a search

A knowledge object is a way to organize a knowledge base so that different instructional algorithms can use the same knowledge objects to teach the same subject matter content. The Knowledge object model is slightly different than other feature areas in that we expose a set of abstract Sales force objects that are not directly used when you create articles. These abstract objects then contain concrete derivations that you'll actually use when creating articles.

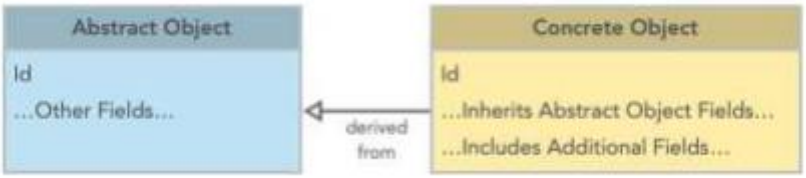

Figure 1: Abstract Object and Concrete Object

The core idea of the use of learning objects is characterized by the following: discoverability, reusability, and interoperability. To support discoverability, learning objects are described by Learning Object Metadata, formalized as IEEE 1484.12 Learning object metadata. To support reusability, the IMS Consortium proposed a series of specifications such as the IMS Content package. And to support interoperability, the U.S. military's Advanced Distributed Learning organization created the Sharable Content Object Reference Model.

Learning objects were designed in order to reduce the cost of learning, standardize learning content, and to enable the use and reuse of learning content by learning management systems E-Learning is a platform to learn anywhere, anytime and through any device and the components of e-learning are content design, development, management, and content delivery. It also includes Metadata

standards. The learning content and its design focuses on learning resources like Learning Objects (LOs).

LITERATURE REVIEW

Ariyajunya, B., Chen, Y., Chen, V. C. P. & Kim, S. B. (2020) Data mining is the extraction of projecting information from large data sets, is a great innovative technology. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Web sites contain millions of unprocessed raw data. By analyzing this data new knowledge can be gained. Since this data is dynamic and unstructured traditional data mining techniques will not be appropriate. In this paper we discuss about data mining techniques. In this paper a survey of the works done in the field of social network analysis and this paper also concentrates on the future trends in research on social network analysis. This paper presents study about social networks using Web mining techniques.

Chan, S. W. K. & Chong, M. W. C. (2020) along with the technology of big data analytics being mature gradually, the concept of business intelligence arose through the whole economic globe. Enterprises tend to explore all kinds of deeply hidden features from mass information networks, with prevalent and intelligent analytics skills such as data mining, process mining, web mining or text mining. Business corporations aim to make predictive decision or take action towards a targeted goal with help of this mined valuable information. Social media is one good example of this. Marketing seems the most direct beneficial field of social media analysis. There has been a number of literature involving the role of social media in marketing and the relationship between market variables and online users generated behaviors Likewise, Senadheera affirmed that social media could act as an information system which encourage enterprises to adopt more business applications via social media platforms, for example interact with lower cost with their customer for feedback.

Chen, T., Xu, R. F., He, Y. L. & Wang, X. (2019) Machine learning is probably the better known and most familiar technique in the field of data mining. It originated from computer gaming and artificial intelligence that give the computer the ability to learn without being explicitly programmed Machine learning tasks are generally classified into two situations, supervised learning and unsupervised learning. Supervised learning involved training computer with presented examples. The goal is to learn the rules from the training set and then apply them into new test task. Unsupervised learning techniques do not require labeled data leaving its own to explore patterns or structures. A very good example to illustrate this is AlphaGo Zero which is an artificial-intelligence chess program from Google Deep Mind team. It mastered the chess game without any human data or guidance, which is strongly beating the previous version which did require a large training database to teach

Bandaru et al., (2017) Machine learning system could realize different function, such as classification, prediction and explore association. Different data mining techniques or algorithms (i.e support vector machine, decision tree) could

be employed in working systems to achieve the expected function with evaluation focusing on accuracy rate lists the sample papers on how to use different tools to build data mining system.

Choi, Y. & Lee, H. (2020) this is typically considered as supervised classification methods and Traore used this method to process a set of satellite images, classifying new areas into epidemic risk or not an epidemic risk in order to find the link between the geographical area of the Niger River and the spread of the cholera epidemic.

Carneiroa et al. (2017) by the latest addressed the problem of fraud detection in order to build a risk scoring system which will predict a Fraud suspicion score for each order. SVM, logistic regression and random forests are employed and compared in this experiment. The latter two data mining techniques are not discussed in this paper due to its non-universality. Prediction is fundamentally based on the relationship between a thing that you can know and a thing you need to predict. A classification problem could be seen as a predictor of class as well. Garcia-Rudolph and Gibert used decision tree algorithm which is a classification algorithm to build data-driven models for Neuro Rehabilitation Ranger which could support the therapist for assigning the most appropriate cognitive rehabilitation plan to each patient.

Deng et al., (2017) Semantic orientation approach(SOA) is more common approach which uses predefined sentiment lexicons such as Word net, Senti Word Net providing lists of sentiment words, or uses external data sources such as corpus providing massive text data sentiment expression and dictionary showing the polarity of words By counting the words from tweets or messages that match categories in lexicon, this approach recognizes words with positive polarity, negative polarity or no polarity (neutral). Stated that statistical analysis and pattern matching could be applied to match words from text documents with the sentiment lexicon Scores are given to matched words then the document-level sentiment scores are calculated from both positive and negative score

Daniel et al. (2017) in the field of financial market utilized 4 text analysis tool based on lexicons to assess sentiment of tweets from financial community which gave a basis for event popularity detection on financial market. Examined the negation scope of financial news using net-optimism which measured the content according to the frequencies of words from the pre-defined dictionaries in the domain of social media research, Stieglitz and Dang-Xuan examined whether sentiment from social media was associated with people’s information sharing behavior using a lexicon to extract sentiment level Chung and Zeng described a system of imood which addressed the sentiment and network analysis based on lexicon orientation approach in order to build a framework for social-media-based public policy informatics. More interestingly, a recent paper by Deng proposed a lexicon expansion method to adapt existing sentiment lexicons for domain-specific sentiment classification.

ISSN: 2233-7857 IJFGCN

LEARNING OBJECT (LO)

LO according to experts is "a reusable, media-independent chunk of information used as a modular building block for e-learning content". It is "any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning" These objects have metadata which refers to description that facilitates & administrates these objects. The various metadata standards are Dublin Core Metadata, IEEE Learning Object Metadata, IMS Global Learning Consortium, and Advance Distributed Learning Initiative (ADLI). The IEEE metadata standard is widely used and it aims to develop accredited technical standards, and guides learning technology. It is grouped under categories like general, life cycle, meta-metadata, educational, technical, rights, relation, annotation, and classification categories. The other providers like IMS content packaging model defines a set of structures used to exchange the learning content.

The content of an instructional unit is referred as LOs. These are reusable, independent small pieces of information used as building blocks for e-learning content. Many structures and features like Learning Objective, Metadata and goals are proposed for LOs. The other key features an LO should have are reusability, interoperability, durability and accessibility. A LO may be a small text, a map, a web page, audio, video or any element that may be used within a course.

Reuse of the LO enables an LO to be used in different ways and in different programmes, and is referred to as "any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning" (IEEE, LTSC). They are rich, interactive material, and are based on granularity; usually discusses a particular subject or learning outcome. Special tools are required for the creation of it and are followed by a review process.

DATA SOURCES IN LEARNING OBJECTS

The massive number of data produced from the user interactions with repositories attracts researchers, educators, decision-makers, and many more stakeholders. Accordingly, a vast number of learning objects and resources are distributed across different datasets and repositories using several metadata standards and access techniques. These learning objects repositories (LOR) have spread widely in an attempt to satisfy the need to learn anywhere and anytime.

Data are produced and captured from user interactions with repositories. There is a great potential in this kind of data, concretely in its processing and analysis as it can inform decisions about future learning patterns and teaching practices. Analytics is a multidimensional disciplinary field, as it applies different techniques to gain understanding from raw data.

Learning objects is one of the major instructional concepts that have driven resource creation due to its potential for reusability, generatively, adaptability, and scalability There are many considerable efforts tried to develop a common definition of learning object which is complex, but still there is a need for a widely acceptable definition that would serve as a framework for further developments.

ISSN: 2233-7857 IJFGCN

Some of those definitions are presented in. In spite of the fact that there is no unified definition for the learning object, most of the researchers agreed on some functional requirements needed to be considered while creating learning objects. In the below subsections, reusability concept, the standards, and metadata will be discussed in more detail and how it affects learning object repositories will be demonstrated.

1. Current Approaches in Evaluating the Quality of Learning Objects

The emergence of huge number of learning objects repositories urged the need for evaluation tools and models to assess the effectiveness of the varied learning objects. Many researchers found it very challenging to create a systematic tool to evaluate the quality of learning objects especially with the multiple technical, functional, and pedagogical requirements needed to be addressed in developing LOs as well as the variety of their metadata and quality metrics. Many evaluation rubrics, tools, and models have been proposed in different papers in order to facilitate the evaluation process. Some of them are mentioned below: proposed the Learning Object Review Instrument (LORI) which is an evaluation framework designed to support collaborative critique of multimedia learning resources. LORI (version 1.5) is a rubric that measures nine critical dimensions of quality spanning pedagogical concerns, technological issues, and user experience factors: Evaluators of the learning objects should provide ratings for each dimension from 1 to 5 scales. The results revealed four distinct constructs: interactivity, design, engagement, and reusability which showed acceptable internal and inter-rater reliability. The study demonstrated that the four constructs were significantly and positively correlated with student learning performance. MERLOT is an online repository that contains reference materials, interactive programs, quizzes, lecture notes, and simulations used to enhance the college learning level. Its resources are classified into nine discipline areas: Arts, Business, Education, Engineering, Health and Human Sciences, Humanities, Math and Statistics, Science and Technology, and Social Sciences. Each major area is divided into multiple sub discipline categories.

2. Learning Object Repositories (LOR)

LORs help users to search for learning material and they provide simple and advanced searches. Simple search results are given based on input keywords. The advanced search allows users to specify values for specific metadata elements and filters the learning material according to the need of the user. Federated search provides the facility of searching learning materials from other repositories. The well-known LORs are MERLOT, EDNA, CGIAR, and CAREO, HEAL. Most of these LOR follow IEEE-LOM metadata standard and metadata annotation is done manually

The massive number of data produced from the user interactions with repositories attracts researchers, educators, decision-makers, and many more stakeholders. Accordingly, a vast number of learning objects and resources are distributed across different datasets and repositories using several metadata

standards and access techniques. These learning objects repositories (LOR) have spread widely in an attempt to satisfy the need to learn anywhere and anytime.

Learning object repositories (LORs) refer to sort of digital libraries or electronic databases where educators can share, use, and reuse different learning objects (LOs) in order to build online learning opportunities for their students. The repository can consist of one database or more linked together using a search engine. LORs can be distinguished based on their infrastructure in four types

- Centralized LOs and centralized metadata: This type locates the learning objects and their metadata on a centralized server. Learning objects retrieval in such infrastructures depends on the server's empowerment and effectiveness. In addition, this infrastructure lacks the flexibility to grow.

- Centralized LOs and distributed metadata: It locates the learning objects in a centralized server, while puts the metadata away from LOs to minimize the problem of having excessive processing costs on centralized servers. By applying this infrastructure, different metadata standards will be applied to learning objects referencing as well as it might be difficult to keep track of all metadata locations.

- Distributed LOs and centralized metadata: These types keep the indexing and metadata centralized in a server and provide links to LOs stored in distributed repositories over the Internet. This type is the most effective one and the most commonly used. This strategy may help in reducing costs because it is not necessary to rely on a robust server infrastructure. On the other hand, the probability of depending on a centralized indexing server is very high, which means any failure will interrupt the accessing of LOs repositories.

- Distributed LOs and distributed metadata: This is a fully distributed architecture. It connects multiple indexing servers with multiple LO databases. This architecture can be done completely in a peer-to-peer way without the need to a centralized server.

Learning Object Repositories (LORPs): where LOs reside within a database on the same server that hosts the web-enabled gateway to the collection

- Learning Object Referatories (LORFs): store only the indexing and metadata and provide links to LOs distributed over the web.

- Open Courseware Initiatives (OCW): provide free and open digital publications organized as courses which can be reused in a learning setting. However, OCW sites do not identify themselves as learning object repositories, even if they comply with the definition given at the beginning.

- Learning Management Systems (LMS): also store a great amount of learning material that is shared in a small community of the teacher and the

students of a course. Even if they are not open, these systems can also be considered as LORs. In Table 5, there are some examples of LORs

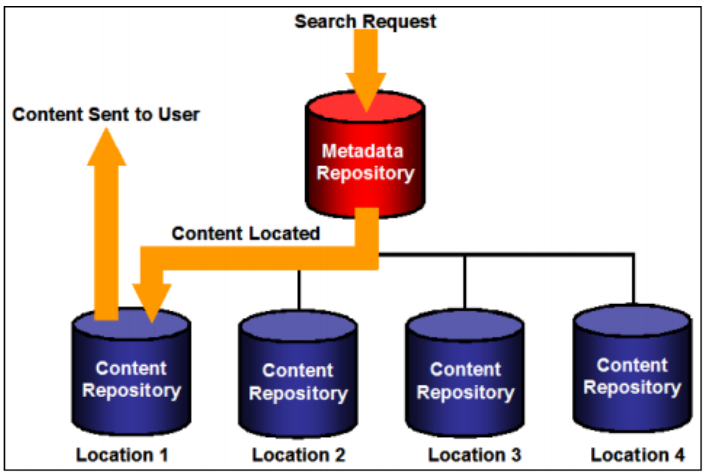

Figure 2: Learning Object Repository – Centralized Metadata

LEARNING OBJECT METADATA BASED REPOSITORIES

Learning objects are the content components that are meant to be reusable in different contexts. These learning objects are associated with metadata, so that they can easily be searchable and manageable. As the international standardization in this area is making a fast progress, the number of learning object repositories are also growing rapidly. A LOM repository or learning object repository stores both learning objects (LOs') and their metadata.

A learning object repository allows users to search and retrieve learning materials from the repository. It supports simple and advanced search, as well as browsing through the materials. In simple search, it returns the search results against the given input keywords. A learner needs to search specific learning materials according to his requirements. The advanced search allows users to specify values for specific metadata elements to filter learning materials to meet the user's specific need. Browsing allows the users to descend in a tree of disciplines and sub-disciplines to access the learning objects available in the repository. Here we will discuss the features and characteristics of some of the existing learning object repositories.

ARIADNE the European digital library project was initiated in 1996 by European commission's telematics for education and training program. Since then, an infrastructure has been developed in Belgium and Switzerland for the production of reusable learning content, including its description, distributed storage, and discovery, as well as its exploitation in structured courses. The core of this infrastructure is a distributed library of digital, reusable educational components called the Knowledge Pool System (KPS). It is actively used in both academic and corporate contexts. The KPS content (Duval and Hodgins is oriented more toward technical science, strongly represented by computer science,

economics, electronics, health science, transportation and life science. The KPS is a reference library.

The KPS includes descriptions (metadata), as well as the documents themselves, making it easier to replicate documents across all nodes of the system, ensuring convenient access without excessive download times. A comparative study of learning Object Metadata, Learning Material Repositories, Metadata Annotation & an Automatic Metadata Annotation Tool 109 the ARIADNE includes a set of metadata from general, technical and educational categories. The ARIADNE includes the traditional metadata title, author and publication date, which are generally used in a library.

THE LEARNING OBJECT MANAGEMENT SYSTEM AGORA

The framework AGORA (from a Spanish acronym that means Help for the Management of Reusable Learning Objects) aims to provide an infrastructure that supports the development of Instructional Design activity. Particularly it provides solutions for Learning Objects management. As result of the project, an operational version of the system has been designed, implemented and tested. The main components of AGORA are:

- A knowledge base consisting of instructional ontologies. The initial version includes models of Instructional Design and methodology for the population, editing and refinement of Instructional Engineering ontologies.

- A system for automatic discovery of knowledge about Instructional Design based on KDD techniques.

- A module for Learning Objects management. It aims to cataloguing, development and processing Learning Objects based on accepted international standards. It includes mechanisms for automatic metadata generation and a repository management system.

- A semantic metasearch engine specialized in e-learning resources available in multiple repositories through a semantic approach that improves the chances to get relevant results according to the user´s instructional needs.

- A method and a tool for assessing the quality of digital learning resources

- A recommender system capable of supporting LOs design, search, recovery and reuse activities, based on the resource development requirements and the teachers profile´s.

CONCLUSION

The purpose of the first research phase was to gain the consensus of a panel of Computer Information Systems (CIS) experts about the functionality for a SOA Data Movement prototype that would resolve the interoperability gap that exists between LMSs and LORs. The Delphi method was leveraged to facilitate the

ISSN: 2233-7857 IJFGCN

process to reach a consensus of CIS experts. The purpose of the second research phase was the development of an Internet- based distributed systems software prototype that served as a proof-of-concept for the use of SOA and the functionality that was agreed upon by the panel of CIS experts. The second phase used the COMET design methodology for the development process of a software prototype to validate the use cases and functional requirements provided by the panel of CIS experts.

REFERENCES

1. Ariyajunya, B., Chen, Y., Chen, V. C. P. & Kim, S. B. (2020). Data mining for state space orthogonalization in adaptive dynamic programming. EXPERT SYSTEMS WITH APPLICATIONS, 76, 49-58.
2. Bandaru, S., Ng, A. H. C., Deb, K., Institutionen för, i., Högskolan i, S. & Forskningscentrum för Virtuella, s. (2017). Data mining methods for knowledge discovery in multi-objective optimization: Part A – Survey. EXPERT SYSTEMS WITH APPLICATIONS, 70, 139-159.
3. Carneiroa, N., Figueira, G. & Costa, M. (2017). A data mining based system for credit-card fraud detection in e-tail. DECISION SUPPORT SYSTEMS, 95, 91-101.
4. Chan, S. W. K. & Chong, M. W. C. (2020). Sentiment analysis in financial texts. DECISION SUPPORT SYSTEMS, 94, 53-64.
5. Chen, T., Xu, R. F., He, Y. L. & Wang, X. (2019). Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. EXPERT SYSTEMS WITH APPLICATIONS, 72, 221-230.
6. Choi, Y. & Lee, H. (2020). Data properties and the performance of sentiment classification for electronic commerce applications. Information Systems Frontiers, 19, 993-1012.
7. Daniel, M., Neves, R. F. & Horta, N. (2017). Company event popularity for financial markets using Twitter and sentiment analysis. EXPERT SYSTEMS WITH APPLICATIONS, 71, 111-124.
8. Deng, S. Y., Sinha, A. P. & Zhao, H. M. (2017). Adapting sentiment lexicons to domain-specific social media texts. DECISION SUPPORT SYSTEMS, 94, 65-76.
9. Giatsoglou, M., Vozalis, M. G., Diamantaras, K., Vakali, A., Sarigiannidis, G. & Chatzisavvas, K. C. (2019). Sentiment analysis leveraging emotions and word embeddings. EXPERT SYSTEMS WITH APPLICATIONS, 69, 214-224.

ISSN: 2233-7857 IJFGCN