International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org Volume 3 Issue 7 ǁ July 2014 ǁ PP.32-35

An Approach Towards Adoption of Pachinko Allocation Model for Automatic Learning Objects Construction And Annotation.

[1] Ajith Jubilson E, [2]Joe Prathap PM, [3]Vimal khanna.V, [4]Jeyabalaji.C,
[5] Dhanavanthini P [6] Vinil Dani W
[1-4] RMD Engineering College Chennai,
[5] Sri Manakula Vinayagar Engineering College
[6] St Xavier's Catholic College of Engineering

ABSTRACT : *Manual construction of learning resources is very expensive since lot of human work is required and only a trained person can create learning objects without flaws. Automatic creation and annotation of learning objects provides effective solution for creation of reusable learning contents. Supervised approach is one which gets the desired output from the user and works towards it. But in unsupervised learning the machine simply receives inputs but obtains neither supervised target outputs, nor rewards from its environment. The existing models Latent Dirichlet Allocation and Hierarchical Latent Dirichlet Allocation does not find correlation between topics also it does not provide options to add new topics.This work proposes an unsupervised technique Pachinko Allocation Model with domain and pedagogy ontology for text segmentation to automatically create and annotate learning objects. The use of Pachinko Allocation Model reduces correlations between topics. In the developing technological world learning resources get increased day by day. Pachinko Allocation Model accepts growing data collection. The enhancement in annotation of Metadata makes the content more flexible.*

KEYWORDS: *Learning objects, E-learning, Ontology, and Learning object metadata.*

I. INTRODUCTION

Technologies and their digital content (E-content) are multiplying by leaps and bounds so we are in need of a model to manage them and for easy retrieval. E-learning is to integrate technologies and to allow learners to learn new knowledge at anytime and anywhere without constraints of time and space. E-learning is an important area where applicability ranges from distance learning, supplementing conventional based learning to need based training for industry. Learning without human intervention is made possible with E-learning. E-content is digital information delivered over network-based electronic devices. The dynamic nature of e-content provides e-learning this facility to adapt to changing learner needs. **(LjiljanaStojanovic, 2001)**

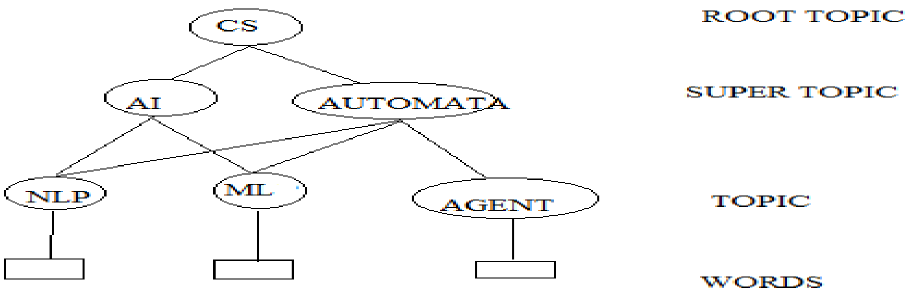
Text segmentation is the process of dividing text into meaningful blocks. It is the process of dividing text according to its context. Manual construction of learning resources is very expensive since lot of human work is required and only a trained person can create learning objects without flaws. Automatic creation and annotation of learning objects provides effective solution for creation of reusable learning contents. This process of text segmentation is important in the e-learning scenario in order to provide relevant segmented content as required by the learner. If proper segmentation is not performed the learner has to read the entire document to learn about a particular topic. Text segmentation is an important component of any language processing tasks. Segmentation of e-learning materials according to topic is an important task in e-learning systems for information extraction The ability to segment documents based on topic enables learners to focus on specific topic rather than to access and analysis whole documents. Longmire states the reason designers might want add learning object capability in their design is that their content gets "more addition in terms of value" which in many cases will prove to be many times effective on the grounds of most critical design factors like costs, development time, and learning effectiveness.

II. PACHINKO ALLOCATION MODEL

The Pachinko Allocation Model structure comprises of an arbitrary DAG, in which each leaf node is leagued with a word in the vocabulary, and each non-leaf internal node relates to a topic, having a distribution over its children. An internal node whose children are all leaves would relate to a conventional topic. But some internal nodes may also have children that are other topics, thus maps a mixture over topics. With various nodes, PAM finds not only correlations amongst words, but also correlations among topics themselves. PAM with the

DAG structure is extremely flexible. DAG is a simple hierarchical tree, or an arbitrary DAG, with cross-connected edges, and edges skipping levels. The nodes can be completely or sparsely connected. The structure could be fixed previously or cultured from the data. PAM provides a general basis for which several prevailing models can be viewed as special cases. PAM provides a powerful means to describe inter-topic correlations and extract large numbers of fine-grained topics. It also presents an ultimatum to determine the appropriate DAG size and structure for a particular dataset.

Figure 1. Four Level Pachinko Allocation Model



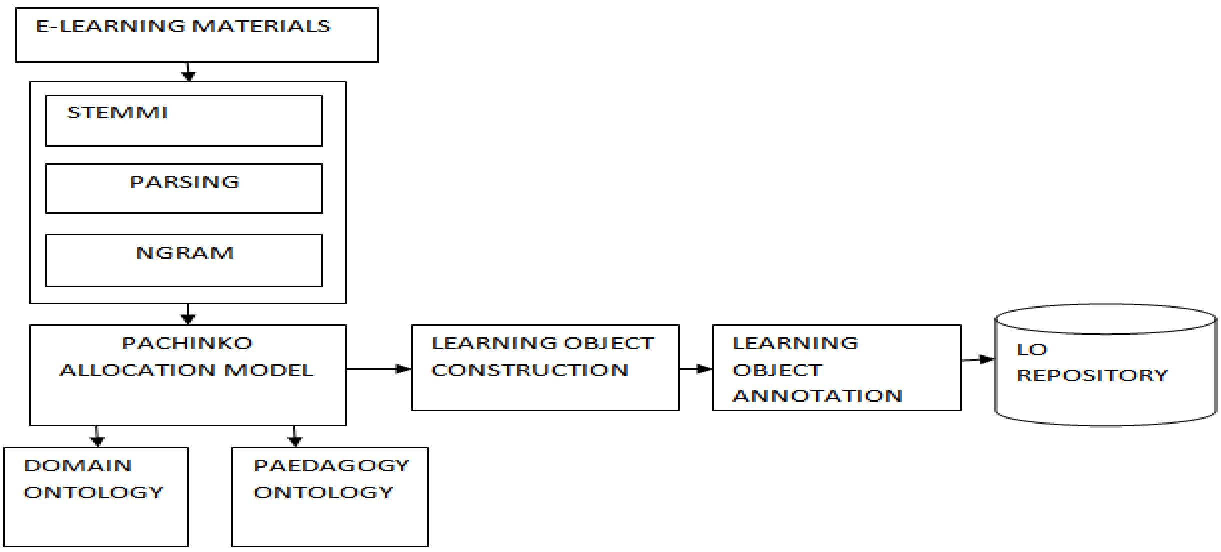
Pachinko Allocation Model is an extension of Latent Dirichlet Allocation and hierarchical Latent Dirichlet Allocation. Here we focus on discovering topics organized into hierarchies

A. **LATENT DIRICHLET ALLOCATION :** The standard LDA topic model represents each document as a mixture of topics. Documents in LDA are linked only through a single non-informative Dirichlet prior. The model therefore makes no attempt to justify the distribution of topic mixtures. Latent Dirichlet allocation (LDA) is a generative probabilistic model of a corpus. The main motive is that documents are represented as arbitary mixtures over latent topics, where each topic is characterized by a distribution over words.

B. **HIERACHICAL LATENT DRICHLET ALLOCATION :** The hLDA model represents the distribution of topics within documents by organizing the topics into a tree. Each document is produced by the topics along a single path of this tree. When understanding the model from data, the sampler substitutes between choosing a new path through the tree for each document and assigning each word in each document to a topic along the chosen path. In hLDA, the value of the distribution of topic mixtures depends on the quality of the topic tree. The assembly of the tree is learned along with the topics themselves using a nested Chinese restaurant process. The usage of this hierarchical LDA model makes the segmentation process flexible to accommodate the growth of large e-learning materials.

Figure 2. Architecture Diagram of the proposed work automatic LO construction using

PAM



III. AUTOMATIC LEARNING OBJECT CONSTRUCTION

Creation of content in the form of standardized learning objects is important in the e-learning scenario since they can be flexibly assembled to deliver courses with different learning objectives with different learning levels and skills. One of the major bottlenecks in the adoption of e-learning is the time and labor required in the manual construction of learning objects. Automatic construction and annotation of learning objects reduces e-content authoring and development time by enabling reuse of available content tailored for different learning purposes. Automated annotation of learning objects with standardized metadata annotation schemes provides universal discoverability and interoperability to search and reuse learning resources

Automatic metadata annotation of each learning object is done according to IEEE LOM standardizing the segmented domain and pedagogical topics. The segmented topics of each segment obtained through the segmentation process contain the domain topic and pedagogical role according to the ACM classification ontology and pedagogical ontology.

IV. LEARNING OBJECT REPOSITORY

Group of learning objects collectively forms the Learning Object Repository (LOR). It is important to annotate the learning objects in the repository appropriately. It must carefully select the set of most relevant attributes that will be used as metadata for the learning objects. However, in order to facilitate the sharing and reuse of learning objects across different information repositories or learning management systems, it is recommended that the learning objects should be associated with some common metadata standard. The IEEE LOM standard is used to create the Metadata. Many contributors find the task of manual annotation and assigning of meta-tags monotonous, and sometimes the tagging is not done agreeably. The development of a repository with manually annotated learning materials is expensive in terms of the time and effort required. So the automated annotation is done which reduces the effort and has more accuracy.

V. CONCLUSION

Pachinko Allocation Model is used to automatically create learning object with the aid of domain and pedagogy ontology to produce quality Learning objects. Since Pachinko Allocation Model uses Directed Acyclic Graph it accepts growing data collection. PAM provides a powerful means to describe inter-topic correlations and extract large numbers of fine-grained topics it also presents a challenge to determine the appropriate DAG size and structure for a particular dataset.

In this work a model for Pachinko Allocation Model using domain ontology was designed for automatic learning objects construction from e-learning materials. The likelihood values are computed for various inputs in Pachinko Allocation Model and are compared with the existing graphical models. P_k metric which is an error metric which finds the segmentation quality. The learning object repository is created along with the learning object metadata. Further the following fields of LOM standard such as Semantic Density, Difficulty and copyright are annotated to make the Learning object more qualitative.

In future the Metadata can be extended for role, identifier, interactivity type, taxon path, source etc. to make the repository more usable. The Pachinko Allocation can also be used for image and video apart from learning objects.

REFERENCES

[1] Warren Longmire,"A Primer on Learning Objects", Available at

[2] http://www.astd.org/LC/2000/0300_longmire.htm.

[3] LjiljanaStojanovic, Steffen Staab, Rudi Studer,"eLearning based on the SemanticWeb",in Proceedings of theWorld Conference on the WWW and Internet University of Karlsruhe, Germany,2001.

[4] EliB. Cohen, MalgorzataNycz,"Learning Objects and E-Learning an Informing Science Perspective", in Proceedings of Interdisciplinary Journal of Knowledge and Learning Objects, 2006.

[5] AndreasHolzinger, Josef Smolle, Gilbert Reibnegger,"Learning Objects (LO): an Object-Oriented Approach to Manage e-Learning Content", in Handbook of Research on Informatics, 2006.

[6] Balatsoukas, P., Morris, A., & O'Brien, A," Learning Objects Update: Review and Critical Approach to Content Aggregation",in the proceedings ofEducational Technology & Society,2008.

[7] McGreal, R Online Education Using Learning Objects RoutledgeFalmer: London and New York Hardback, ISBN: 0 4153 3512 4 2005.

[8] Yee WhyeTeh, "Dirichlet Process", in Proceedings in University College London, 1995.

[9] Saturnino Luz,"Machine Learning for NLP: Unsupervisedlearning techniques",in the Proceedings of ESSLLI'07 Dept. of Computer Science, Trinity College Dublin, Ireland, 2007.

[10] David M. Blei, Andrew Y. Ng, Michael I. Jordan, "Latent Dirichlet Allocation", in Proceedings of Journal of Machine Learning Research , pp.993-1022, 2003.

[11] Paramsjeet Singh Saini, Marco Ronchetti, Diego Sona, "Automatic Generation of Metadata for Learning Objects", in Proceedings of Advanced Learning Technologies,Sixth International Conference, 2006.

[12] Kun Hua Tsai, Ti Kai Chiu, Ming Che Lee, Tzone I Wang A "Learning Objects Recommendation Model based on the Preference and Ontological Approaches",in Proceedings in National Chung Kung University, Tainan, Taiwan,2006.

[13] AmalZouaq, DraganGasevic, MarekHatala b, "Towards open ontology learning and filtering", in the proceedings of Information Systems Volume 36, Issue 7, pp. 1064-1081, November 2011.
[14] Elias ZavitsanosGeorgiosPalioura sGeorge A.Vouros, "Discovering Subsumption Hierarchies of Ontology Concepts from Text Corpora", in Proceedings ofIEEE/WIC/ACM International Conference on Web Intelligence IEEE Computer Society Washington, DC, USA, 2007.
[15] Adler Perott, "Hierarchically supervised latent Dirichlet allocation (HSLDA), a model for hierarchically and multiply labeled bag-of-word data", in Proceedings of Information Systems, pp. 1064-1081, November 2010.
[16] K.Sathiyamurthy, T.V.Geetha, "Topic Segmentation and Evaluation Measures for e-learning Based on Domain and Pedagogy Ontology", in International Journal of Computer Applications, Vol 26– No.6, July 2011.
[17] Adler Perotte Nicholas Bartlett Noemie Elhadad Frank Wood "Hierarchically Supervised Latent Dirichlet Allocation", In Proceedings of 2nd Workshop on human motion Understanding, Modeling, Capture and Animation, 2007.
[18] Wei Li, Andrew McCallum, "Pachinko Allocation: DAG-Structured Mixture Models of Topic Correlations", in the Proceedings of the 23 rd International Conference on Machine Learning, Pittsburgh, PA, 2006.