Middle-East Journal of Scientific Research 23 (9): 2074-2080, 2015
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2015.23.09.22426

Certain Investigation of Identify the New Rules and Accuracy Using SVM Algorithm

[1]*S. Yogadinesh,* [1]*S. Oswalt Manoj,* [1]*K. Srihari,* [1]*S. Rajesh,*
[1]*R.N. Devendran,* [2]*K. Ayyaparaja and* [1]*E. Udayakumar*

[1]Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, India
[2]Department of CSE, Theni Kammavar Sangam College of Technology, India

Abstract: Nowadays, technology-supported learning systems (TSLSs), such as intelligent tutoring systems (ITSs), adaptive hypermedia systems (AHSs) and, especially, learning management systems(LMSs) such as Moodle or Blackboard, are being widely used in many academic institutions and becoming essential for education. The Domain Module is considered the core of any TSLSs as it represents the knowledge about a subject matter to be communicated to the learner. In the existing system, proposed a DOM-Sortze it is a system that uses (NLP) natural language processing techniques, heuristic reasoning andontology is the semiautomatic construction of the domain module initial electronic textbooks. But in this system, still lack in the identification of pedagogical relationships. This is needed to improve in this system. In other words, DOM-Sortze system is not able to including the new rules of the pedagogical relationships. To overcome this issue, we are using learning techniques to learn the new rules in the pedagogical relationships. In our proposed system, we are proposing the SVM (support vector machine) learning approach intended for learning process. Our machine learning methods are used to infer new rules in order to improve the identification of pedagogical relationships or the DRs in the electronic textbooks.

Key words: Heuristic Reasoning · Ontology · Support Vector Machine

INTRODUCTION

Ontology learning is nothing that the gathering of domain with helps of different resource which gathering in the automatic or semiautomatic ways which are N number of works [1]. Automatic in the sense it will gather the domain ontology from different resource by automatically in semiautomatic process gathering of domain is integrated by manual process [2]. This ontology process combine the NLP technic machine learning technic to build the domain ontology are it is used to enhance the domain ontology. Pedagogical relationships is a structural way it is four process are in LDO Gathering is A and part of or Prerequisite and next [3] [4]. The X ^{isA}Y X relationships is declares topic is a particular mild of Y. Part of X$^{part\ Of}$Y which may differ from X and Y. The Prerequisite is Y ^{Prerequisite} X if we attempt to learn the topic Y we should know the master state relationship of topic X.while X ^{next} Y precise it is suggest to learn topic Y rights after control topic Y [6].

Outline Analysis: Read the overview of topic that document are structure way it used to semiautomatic so they are given to summarized details and given meaning information and get low cost .they are two outline analysis Basic analysis and Heuristic analysis [7] [8].

Basic Analysis: When mining an index item it surfs for sub item and its pedagogical relationships in the internal representation in order to learn its sequence the index are sub item is relations are [9] [10] isA and part of.

Heuristic Analysis: Checks whether the recognized relationships match with in define heuristic to be unique [11]. They are two relationships in heuristic it is structural in isA and part Of relationships and sequential relationships in Prerequisite and next.

Heuristic for Structural Relationships

Multi Word Heuristic: The way used to extract the information about isA relation. It checks isA relationship is Plausible probable or not via taxonomic relationships.

Corresponding Author: S. Yogadinesh, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, India.

Entity Name Heuristic: The relationship that exists between the item and sub item of entity isA relationships [12].

Acronym Heuristic: The most frequently used words are name or listed under acronyms.

He-MWH: Using the He-MWH it has to be check that sub item forms the multiword term of outline item [13].

Possessive Genitives Heuristic: It uses part odf sequence relationship to diagnose the relation between item & sub item for reference.

Keyword Heuristic: The keywords are identified based on the outline analyzed and can be easily if new once are found [14].

Heuristic for Sequential Relationships: Reference Heuristic: if it used the sentence of one words the next word can be automatically generated foe Reference Heuristic [14] [15].

Document Body Analysis: Document body analysis gets improve than the previous phase with new domain topics [16]. The hybrid method combine the NLP technique and statistical method in term of extraction are used in last few years A set of pattern such as ((A|N) +|((A|N) * (NP) ?(A|N)*)N [17] [18] to the set of candidate terms A is adjective N is Noun, P is preposition are used go fitter or get appropriate result applied some term hood measure to rank the set of candidate terms [19]. In DOM-SORTZE, term extraction is done by ErauztermUsual noun phase structure is used for term extraction Basque to get new do main topics .Domain relatedness is determined using different measures which consist of one word or multiword terms gathered by Erauzterm [20].

Identifying New Relationships among Topics: In this Process new structure way is identified from electronic document using a pattern based approach. this pattern recognize relations between domain topics based on syntactic structures found in sentence any domain topics appearance is done to label the internal representation of document Nested domain topics is identified and proposed [21]. Grammar driven analysis is used to identify a set of sentences which relates two or more domain topics the grammar is set of rules which describes syntactic structure that corresponds to pedagogical relations constraint grammar formalism is used to develop and apply the grammar on documents. It is the most successful syntax analysing and disambiguation system for recognizing structural relationships A and part of and Prerequisite sequential relationship grammar is identified [22] [23]. The rules are defined after empirical analysis of textbooks that corresponds to primary school for grammar definition 13 rules for A relationship 6 part of and 1 prerequisite relationship are used. It include the equivalent for Hearst's pattern are used in many ontology learning approaches to identify the taxonomic relationships hypernyms, hyponyms from syntactic structure like [24] [25] NP0, NP1, NP2…. (and 10r,) NPn, where NPi is noun phrases corresponds to a term or topic.

Hypernyms is noun whose meaningis included in the meaning of another general word Example. In terms of different type of colour hypernyms is the colour that is in range between crimson and violet [26] [27]. Hyponyms is also refers to noun and refers to category of actions. Examples: look can be stare, view gage and peer [28] [29]. It refers the different type of look.

Generation of the DRs: The DRs is identified from relevant text fragments which corresponds to definition example facts, theories and other problem statements for 100 topics [30] [31]. 100 topics are first labelled as document internal representation with part of speech, then DR grammar is used find text fragments which includes set of rules which define the different pattern that have been found in the text book for primary school and tested on a set of textbooks for primary school. The patterns are most common syntactic structure. Constraint grammar is used for developing electronic document from the DRs Grammar [32] [33].

To observe performance, the DR grammar was texted on electronic text books. Initially defined rules are removed from final version of DR grammar [34]. The identified DRs contain the sentence which triggered the rule for another DR which refers to the same topics. Every DR is labelled with domain and rules are identified and it can be used for later LO annotation process. The DR grammar is identified by the DR grammar and enhanced in two ways to make them more accurate. [35] [36] When consecutive DRs similar they are combined and similarity measures are defined. Two consecutive atomic definitions are combined to get more comprehensive DR. To keep the cohesion of the DRs, previous fragments are added to each DR. The comprehensive DR is built as aggregate DR of lower granularity and keeps the information of similarity rates and composed techniques [37] [38]. In every DR the referred topics and the DR grammar rules are kept.

LDO Gathering: By identifying and gathering DRS, Los is generated. For particular educational purpose, consistent fragments of documented which is related to one or more topics [39] [40]. NLP techniques are used in ontology driven process for identification and extraction. In this work LO generating approach is domain independent. Learning session (e.g. definition, etc...) uses document which will be referred by DR, while LO refers to reusable DR. The LO generation process is carried out by Erauz Ont, which is part of the DOM Sortze framework [41].

The LOs gathering from electronic document is done by following task: [42] [43] generating DRs, annotating the DRs to become LOs and finally storing the generated LOs in a LOR for further use. The LDO and ALOCOM ontology which build the LOs from gathered DR and LOs are stored in LOR for reuse [44].

SVM Learning: To identify the pedagogical relationship, machine learning methods are newly introduced in the proposed system. For this purpose SVM as machine learning methods are used. In the training phase new rules in the pedagogical relationship or the DRs in the electronic textbooks are learned and identified. Support Vector Machine is based on decision planes which define decision boundaries. A decision plane separates a set of object from different class memberships. To separate the two classes, SVM modelling algorithm finds optimal hyper plane with the maximal margin to separate two classes, which requires solving the following optimization problem.

Maximize:

$$\sum_{I=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j k(x_i,x_j) \quad (1)$$

Subject to,

$$\sum_{i=1}^{n}\alpha_i y_i = 0 \quad (2)$$

where $0 \le \alpha_i \le b$ i = 1,2,... ... ..,n Where α_i is the weight of training sample x_1. If $\alpha_i > 0$, x1 is called a support vector b is a regulation parameter used to trade-off the training accuracy and the model complexity so that a superior generalization capability can be achieved. K is a kernel function, which is used to measure the similarity between two samples. A popular radial basis function (RBF) kernel functions. This process is repeated *k* times for each subset to obtain the cross validation performance over the whole training dataset. If the training dataset is large, a small subset can be used for cross validation todecrease computing costs. The following algorithm can be used in the classification process.

Input: Sample x to classify training set T = $\{(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)\}$; number of nearest neighbours k.

Output: Decision y_p Î {-1, 1}

Find k sample (x_i, y_i) with minimal values of K (x_i, x_i) – 2 * K (x_i, x)
Train an SVM model on the k selected samples
Classify x using this model, get the result y_p
Return y_p

Proposed Algorithm
Input: Number of the training samples (determined in existing system) with dataset *w* as input data point for SVM classification

Output: Classification result i.e., prediction of the pedagogical relationships result.

Procedure SVM (*w*)
Begin
Begin Initialize C=0 Get input file dataset w for training
Read the number of input training dataset W from original dataset

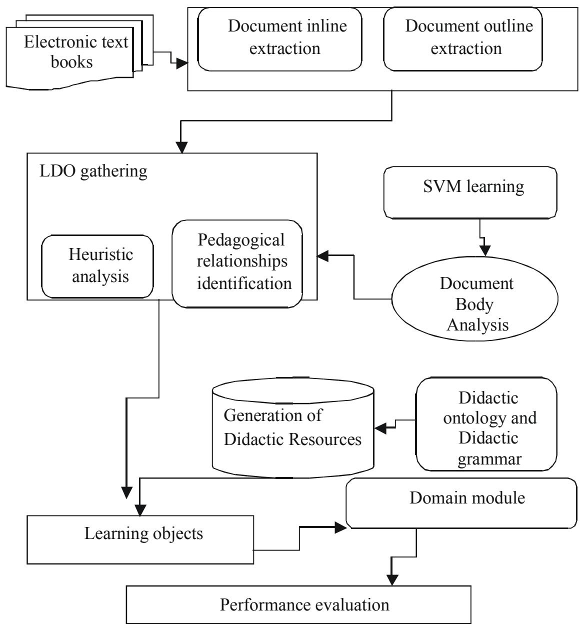
$x_i \cdot w + b = 0$
$x_i \cdot w + b = 1$

Decision function $f(W) = x_i \cdot w - b$

If $f(W) \ge 1$ for x_i is the first class Else $f(W) \le -1$ For x_i is the second class.

The prediction result for (i=1…n) number of training samples $[y_i(x_i \cdot w - b) \ge 1$ Display the result.

Architecture for SVM
Performance Evaluation
Precision: Precision value is calculated is based on the retrieval of information at true positive prediction, false positive. In healthcare data precision is calculated the percentage of positive results returned that are relevant.



$$\text{Precision} = \frac{\textit{True positive}}{\textit{True positive} + \textit{False positive}}$$

Recall: Recall value is calculated based on true positive prediction, false negative. In healthcare, data precision is calculated based on the True Positive Rate. Recall is the part of pertinent instance that are obtained.

$$\text{Recall} = \frac{\textit{True positive}}{\textit{True positive} + \textit{False negative}}$$

TP (True Positive): In a statistical hypothesis test, there are two types of incorrect conclusions that can be drawn. The hypothesis can be inappropriately. A positive test result that accurately reflects the tested-for activity of an analysed. If the outcome from a prediction is p and the actual value is also p, then it is called a true positive (TP);

True positive rate (TPR) =TP/P

P = (TP+FN)

where P is the positive. TP is the True Positive.

TN (True Negative): A result that appears negative when it should not. A true negative (TN) has occurred when both the prediction outcome and the actual value are n is the number of input data.

True negative rate (TNR) =TN/N
N= (TN+FN)

where,
N is the Negative value.
TN is the True Negative.

FP (False Positive): A result that indicates that a given condition is present when it is not. However if the actual value is n then it is said to be a false positive (FP).
False positive rate (α) = FP / (FP + TN)

FN (False Negative): False negative (FN) is when the prediction outcome is n while the actual value is p.
False negative rate (β) =FN / (TP + FN)

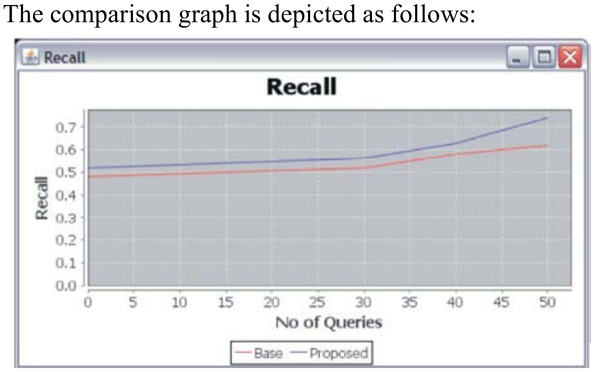
RESULTS AND DISCUSSION

The experimental results were conducted to show the effectiveness of the proposed methodology in terms of precision, recall, Accuracy and F-Measure. This attributes are used to show whether the proposed methodology predicts the domain module correctly or not. The comparison of classification of proposed methodology is done comparing it with the existing method.

RECALL: Recall is the performance measure is used to retrieve result and check whether it is relevant to the document or not. The calculation of the recall value is done as follows:

$$\text{Recall} = \frac{\textit{True positive}}{\textit{True positive} + \textit{False negative}}$$

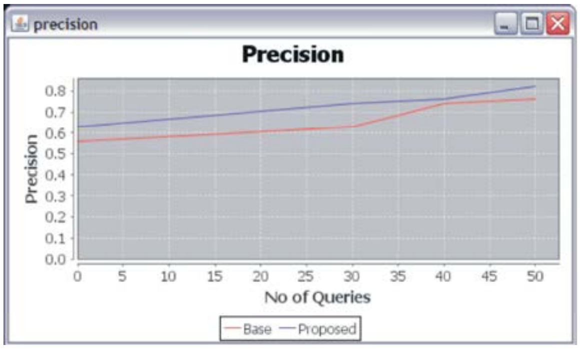
The comparison graph is depicted as follows:



The above graph plots the comparison of the existing methodology against an proposed methodology in terms of the recall measure. In this graph x axis plots the number of queries that are processed and the y axis plots the recall measure value. This graph proves that the proposed methodology provides an better result than the existing methodology.

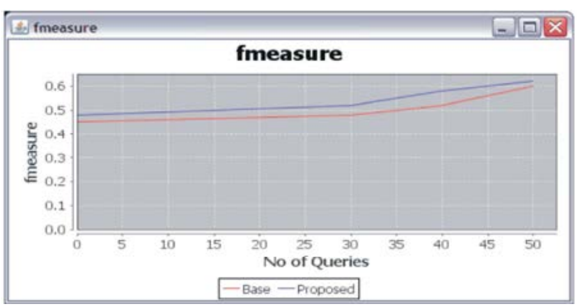
Precision: Precision value is used to indicate that whether the retrieved result is relevant to the corresponding document or not. The precision is calculated as follows:

$$\text{Precision} = \frac{True\ positive}{True\ positive + False\ positive}$$



The above graph plots the comparison of the existing methodology against a proposed methodology in terms of the precision measure. In this graph x axis plots the number of queries that are processed and the y axis plots the precision measure value. This graph proves that the proposed methodology provides a better result than the existing methodology.

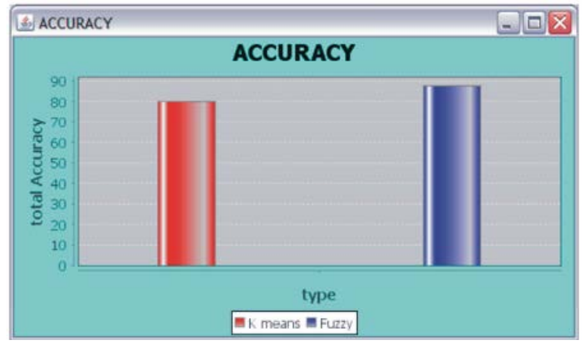
F-Measure $F = 2.\frac{precision.recal}{precision + recal}$



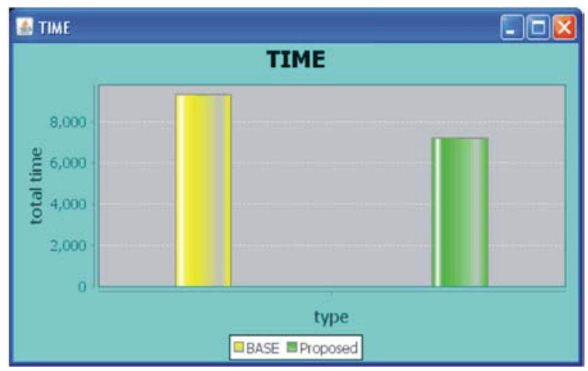
It is a parameter of the test.

$$\text{Accuracy} = \frac{\text{Number of true positives} + \text{number of true negatives}}{\text{Number of true positives} + \text{false positives} + \text{false negatives} + \text{true negatives}}$$

The above graph plots the comparison of the existing methodology against a proposed methodology in terms of the Accuracy. In this graph x axis plots the number of queries that are processed and the y axis plots the F measure value. This graph proves that the proposed methodology provides a better result than the existing methodology.



Time Comparison: In this section total time taken to extract the domain modules by both existing and the proposed methodologies is compared. In the following graph, the comparison is done with the consideration of the accurate domain modules that are extracted. The comparison of time consumption by both existing and the proposed methodology are shown in the following graph.



The above graph plots the comparison of the existing methodology against a proposed methodology in terms of the time consumption. In this graph x axis plots the methodology that are processed and the y axis plots the time measure value in milli second. This graph proves that the proposed methodology provides a better result than the existing methodology.

CONCLUSIONS

In the previous work has presented DOM-Sortze, a system for the semiautomatic generation of the Domain Module from electronic textbooks. The system employs NLP techniques, heuristic reasoning and ontologies for the knowledge acquisition processes. Our proposed system, introduce the machine learning methods which is used for effectively identify the pedagogical relationships. We are using the SVM as machine learning method for this purpose. In this approach, we are learning the new rules in the pedagogical relationships or the DRs in the electronic textbooks in the training phase. Based on the training phase, we can easily identify the pedagogical relationships or the DRs in the electronic textbooks. Experimental results shows that our proposed system is well effective than the existing system. Our system is improves the accuracy of the system, thus undoubtedly the performance of the system is improved.

REFERENCES

1. Aduriz, I., E. Agirre, I. Aldezabal, I. Alegria, O. Ansa, X. Arregi, J.M. Arriola, X. Artola, A.D. De Ilarraza, N. Ezeiza, K. Gojenola, A. Maritxalar, M. Maritxalar, M. Oronoz, K. Sarasola, A. Soroa, R. Urizar and M. Urkia, 1998. "A Framework for the Automatic Processing of Basque," Proc. Language Resources and Evaluation Conf. (LREC '98).
2. Aduriz, I., I. Aldezabal, I. Alegria, X. Artola, N. Ezeiza and R. Urizar, 1996. "Euslem: A Lemmatiser/Tagger for Basque," Proc. EURALEX, 1: 17-26.
3. Agirre, E., O.L. De Lacalle and A. Soroa, 2009. "Knowledge-Based WSD and Specific Domains: Performing Better Than Generic Supervised WSD," Proc. 21st Int'l Joint Conf. Artifical Intelligence (IJCAI '09), pp: 1501-1506.
4. Alegria, I., A. Gurrutxaga, P. Lizaso, X. Saralegi, S. Ugartetxea and R. Urizar, 2004. "An XML-Based Term Extraction Tool for Basque," Proc. Fifth Int'l Conf. Language Resources and Evaluations (LREC '04).
5. Anderson, J.R., 1988. "The Expert Module," Foundations of Intelligent Tutoring Systems, M.C. Polson and J.J. Richardson, eds., pp: 21-54, Lawrence Erlbaum.
6. Bechhofer, S., F. Van Harmelen, J. Hendler, I. Horrocks, D.L. McGuinness, P.F. Patel-Schneider and L.A. Stein, 2004. "OWL Web Ontology Language Reference," http://www.w3.org/ TR/owl-ref/.
7. Buitelaar, P., D. Olejnik and M. Sintek, 2004. "A Prote´ge´ Plug-In for Ontology Extraction from Text Based on Linguistic Analysis," Proc. First European Semantic Web Symp. (ESWS '04), pp: 31-44.
8. Byrd, R.J., N. Calzolari, M.S. Chodorow, J.L. Klavans, M.S. Neff and O.A. Rizk, 1987. "Tools and Methods for Computational Lexicology," Computational Linguistics, 13(3-4): 219-240.
9. Cardinaels, K., M. Meire and E. Duval, 2005. "Automating Metadata Generation: The Simple Indexing Interface," Proc. 14th Int'l Conf. World Wide Web (WWW '05).
10. Chen, W., R. Lu, W. Zhang and H. Du, 1997. "A Tool for Automatic Generation of Multimedia ICAI Systems," Proc. Int'l Conf. Artificial Intelligence in Education (AIED '97), pp: 571-573.
11. Chen P.S.D., A.D. Lambert and K.R. Guidry, 2010. "Engaging Online Learners: The Impact of Web-Based Learning Technology on College Student Engagement," Computers and Education, 54(4): 1222-1232.
12. Cimiano, P. and J. Vo¨ lker, 2005. "Text2Onto—A Framework for Ontology Learning and Data-Driven Change Discovery," Proc. 10th Int'l Conf. Applications of Natural Language to Information Systems (NLDB '05), pp: 227-238.
13. Conde, A., M. Larran˜ aga, I. Calvo, J.A. Elorriaga and A. Arruarte, 2012. "Automating the Authoring of Learning Material in Computer Engineering Education," Proc. 42nd IEEE Frontiers in Education Conf. (FIE '12), pp: 1376-1381.
14. Constraint Grammar: Language-Independent System for Parsing Unrestricted Text," Natural Language Processing, F. Karlsson, A. Voutilainen and J. Heikkila, eds., no. 4, Mouton de Gruyter, 1995.
15. Frantzi, K.T., S. Ananiadou and J. Tsujii, 1998. "The C-Value/NC-Value Method of Automatic Recognition for Multi-Word Terms," Proc. Second European Conf. Research and Advanced Technology for Digital Libraries (ECDL '98), pp: 585-604.
16. Guarino, N., 1997. "Semantic Matching: Formal Ontological Distinctions for Information Organization, Extraction and Integration," Proc. Int'l Summer School on Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology (SCIE '97), pp: 139-170.
17. Hearst, M.A., 1992. "Automatic Acquisition of Hyponyms from Large Text Corpora," Proc. 14th Conf. Computational Linguistics (COLING '92), pp: 539-545.
18. Hughes, T. and D. Ramage, 2007. "Lexical Semantic Relatedness with Random Graph Walks," Proc. EMNLP-CONLL-2007, pp: 581-589.
19. Jones, K.S., 1972. "A Statistical Interpretation of Term Specificity and Its Application in Retrieval," J. Documentation, 60(5): 11- 21.
20. Justeson, J.S. and S.M. Katz, 1995. "Technical Terminology: Some Linguistic Properties and an Algorithm for Identification of Terms in Text," Natural Language Eng., 1(1): 9-27.

21. Kabel, S.C., R. De Hoog, B.J. Wielinga and A. Anjewierden, 2004. "The Added Value of Task and Ontology Based Mark-Up for Information Retrieval," J. Am. Soc. for Information Science and Technology, 55(4): 348-362.
22. Larran˜ aga, M., A. Conde, I. Calvo, A. Arruarte and J.A. Elorriaga, 2012. "Evaluating the Automatic Extraction of Learning Objects from Electronic Textbooks Using Erauzont," Proc. 11th Int'l Conf. Intelligent Tutoring Systems (ITS '12), pp: 655-656.
23. Larran˜ aga, M., I. Calvo, J.A. Elorriaga, A. Arruarte, K. Verbert and E. Duval, 2011. "ErauzOnt: A Framework for Gathering Learning Objects from Electronic Documents," Proc. 11th IEEE Int'l Conf. Advanced Learning Technologies (ICALT '11), pp: 656-658.
24. Larran˜aga, M., U. Rueda, J.A. Elorriaga and A. Arruarte, 2004. "Acquisition of the Domain Structure from Document Indexes Using Heuristic Reasoning," Proc. Seventh Int'l Conf. Intelligent Tutoring Systems (ITS '04), pp: 175-186.
25. Larran˜aga, M., J.A. Elorriaga and A. Arruarte, 2008. "A Heuristic NLP Based Approach for Getting Didactic Resources from Electronic Documents," Proc. European Conf. Technology Enhanced Learning (EC-TEL '08), pp: 197-202.
26. Larran˜aga, M., I. Niebla, U. Ruedat, J.A. Elorriaga and A. Arruarte, 2007. "Towards Collaborative Domain Module Authoring," Proc. Seventh IEEE Int'l Conf. Advanced Learning Technologies (ICALT '07), pp: 814-818.
27. Leidig., T., 2001. "L3-Towards an Open Learning Environment," ACM J. Educational Resources in Computing, 1(1): 5-11.
28. Lentini, M., D. Nardi and A. Simonetta, 2000. "Self-instructive Spreadsheets: An Environment for Automatic Knowledge Acquisition and Tutor Goeneration," Int'l J. Human-Computer Studies, 52(5): 775-803.
29. Liu, B., C.W. Chin and H.T. Ng, 2003. "Mining Topic-Specific Concepts and Definitions on the Web," Proc. 12th Int'l Conf. World Wide Web (WWW), pp: 251-260.
30. Maedche, A. and S. Staab, 2001. "Ontology Learning for the Semantic Web," IEEE Intelligent Systems, 16(2): 72-79.
31. Meire, M., X. Ochoa and E. Duval, 2007. "SAmgI: Automatic Metadata Generation v2.0," Proc. World Conf. Educational Multimedia, Hypermedia and Telecomm. (ED-MEDIA '07), pp: 1195-1204.
32. Morin, E. and C. Jaquemin, 1999. "Projecting Corpus-Based Semantic Links on a Thesaurus," Proc. 37th Ann. Meeting of the Assoc. for Computational Linguistics (ACL '99), pp: 389-396.
33. Ontology Learning from Text: Methods, Applications and Evaluation, Buitelaar P., P. Cimiano and B. Magnini, eds., IOS Press, 2005.
34. Parsad, B. and L. Lewis, 2008. "Distance Education at Degree-Granting Postsecondary Institutions: 2006-07," technical report, Nat'l Center for Education Statistics, Inst. of Education Sciences, US Department of Education.
35. Semi-Automatic Ontology Development: Processes. and Resources, Pazienza M.T. and A. Stellato, eds., IGI Global, 2012.
36. Simon, B., D. Massart, F.V. Assche, S. Ternier, E. Duval, S. Brantner, D. Olmedilla and Z. Miklo´ s, 2005. "A Simple Query Interface for Interoperable Learning Repositories," Proc. 14th Int'l Conf. World Wide Web (WWW '05), pp: 11-18.
37. Ternier, S., D. Massart, F.V. Assche, N. Smith, B. Simon and E. Duval, 2008. "A Simple Publishing Interface for Learning Object Repositories," Proc. World Conf. Educational Multimedia, Hypermedia and Telecomm. (ED-MEDIA '08), pp: 1840-1845.
38. Velardi, P., R. Navigli, A. Cucchiarello and F. Neri, 2005. "Evaluation of OntoLearn, a Methodology for Automatic Learning of Domain Ontologies," Ontology Learning from Text: Methods, Applications and Evaluation, P. Buitelaar, P. Cimiano and B. Magnini, eds., pp: 92-106.
39. Verbert, K., 2008. "An Architecture and Framework for Flexible Reuse of Learning Object Components," PhD dissertation, FaculteitIngenieurswetenschappen, Katholieke Univ. Leuven.
40. Verbert, K., D. Ga_sevi_c, J. Jovanovi_c and E. Duval, 2005. "Ontology- Based Learning Content Repurposing," Proc. 14th Int'l Conf. World Wide Web (WWW '05), pp: 1140-1141.
41. Verbert, K., X. Ochoa and E. Duval, 2008. "The ALOCOM Framework: Towards Scalable Content Reuse," J. Digital Information, 9(1).
42. Vossen, P., 2001. "Extending, Trimming and Fusing Word Net for Technical Documents," Proc. Second Meeting of the North Am. Chapter of the Assoc. for Computational Linguistics (NAACL '01).
43. Word Net: An Electronic Lexical Database, Fellbaum C., ed., MIT Press, 1998.
44. Zouaq, A. and R. Nkambou, 2009. "Evaluating the Generation of Domain Ontologies in the Knowledge Puzzle Project," IEEE Trans. Knowledge and Data Eng., 21(11): 1559-1572.