International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

A Semantic based Terms Relation Research for Classification in Web Information Mining

Sunil Kumar Thota, Tummala Sita Mahalakshmi

ABSTRACT--- The rapid growth of web information and its services in different areas such as e-commerce, healthcare, digital marketing, online booking, etc. is a challenge in providing accurate information in the domain services related to the user's query. The current web information of services classifies the retrieval of the relevant service and assists the classification by supporting the knowledge and classifications of the specific service information. Because of these limitations and the complexity of automatic update mechanisms to see this service information, a large number of non-related service information for a requested query, and getting the required web information of services is a cumbersome problem. This paper proposes Semantic based Terms Relation Approach (STRA) for classifying information for effective classification of WIS on the web. The approach utilize Concept Terms Similarity (CTS) method for the most relevant terms in a service domain and construct a Related Terms Hierarchal Model (RTHM), which will be used for classification. A modified Naive Bayes classifier is used to perform the classification of the web information of services using RTHM, to categorize and present accurately. The experiment evaluation of the proposed approach shows an improvement in the classification of information and achieve a highly related matching results against different number of users queries.

Keywords — Web mining, Domain Services, Semantic Relation, Concept Terms Similarity, Classification.

I. INTRODUCTION

The advent of the Internet and online businesses to support Web Information of services (WIS) is the next level of complexity that provides the flexibility to integrate applications through standard protocols. Classification of services by different mechanisms makes it easier to recognize WIS [1], [2], [3]. They have thousands of communities to make decisions and connect the appropriate category of services, and even WIS are by far the largest interpretation without the need for a unique link that requiring the knowledge of multi-level hierarchical classifications [4], [5]. As a result, a number of services related to a specific customer service request may not be considered during service discovery. It has a full classification of the web service for the need for appropriate expertise and knowledge.

So far, the ontology hierarchy model has emerged as a way to capture knowledge from existing domain, applications, and tasks. It is logical to involve them in the classification process to support semantic applications by providing a domain-oriented visualization [6], [7]. But, with the growth of the semantic web and the existence of many tools to create and manage ontology, the ontology of different domains of knowledge is already available [8], [9]. Their participation in the discovery of knowledge based on automated learning has increased and promising results have been obtained.

The WIS is a fundamental problem in the recommendation and choice of service-oriented computing [9], [10]. Current search in the WIS and recommend either destruction of UDDI records, dominant keyword, or service-based web service policies for QoS [11], [12]. These methods have many limitations and recommended bad performance and are highly dependent on customer input query.

The problem of automatic classification of information and WIS has been widely discussed in the past decade [8], [10], [29], [30]. Demand for supporting tools and methods is increasing as more and more information and services are constantly available in e-libraries, e-commerce, blogs and forums [4], [14] [15], [31] on the other hand, Clear texts regarding the subject matter area are also a very important subject in information processing techniques to discover knowledge. This task can be facilitated by providing means to oversee the classification of the set of texts examined to match domain services [5], [13], [32].

This paper propose a "Semantic based Terms Relation Approach (STRA)" to classifying information for effective classification of WIS. It constructs a "Related Terms Hierarchal Model (RTHM)" utilizing a "Concept Terms Similarity (CTS)" method. The constructed RTHM is used to classify the Web information utilizing a modified Naive Bayes classification method. The novelty of the contribution is to provide regular update of the domain service information through regular updating of the service information in RTHM. This will enhance the much needed web information mining in various domain services categorization and retrieval of the accurate services.

The following paper presents a review of related works in section-2, in section-3 it discuss the proposed Semantic based Terms Relation Approach, in section-4 it present the experimental evaluation and result analysis of the proposal, and section-5 present the conclusion of the work.

Revised Manuscript Received on 30 July 2019.
***** Correspondence Author
Sunil Kumar Thota*, Research Scholar, CSE Department, Gitam University, Vishakapatnam.
Dr. Tummala Sita Mahalakshmi, Professor, CSE Department, Gitam University, Vishakapatnam.

II. REVIEW OF RELATED WORKS

The review of relevant works in literature discusses the interest in achievements in the classification of information services with the support of ontology in [17], [18], [19] the definitions of the argument for grammatical or semantic concepts vary according to classification methods of services [23], [24], [25].

L. Yao et al. [8] proposed an innovative approach that unifies collaborative filtering services and content-based web information of services. The methodology that combines classic collaborative filtering with content-based recommendation uses a three-tier model. Semantic Web Information of Services similarities are considered concurrently with user ratings and provide a mixed approach. This approach is often based on clearly defined Web services for the recommendation limiting the scope of the service recommendation.

H. Dung et al. [13] assessed three major problems with online crawling in terms of careful thinking about the efficiency of information discovery, information coordination, and online information classification. It proposes a new creeping program based on adapting to the solution problem mentioned above. Without a web environment framework, in order to maintain crawler performance, it can crawl learning methods in linguistics. The existence of unmanaged learning vocabulary based on this research lies in the formulation of concepts and metadata closely related to matching the discoveries of the mixed algorithm.

Vogrincic et al. [10] suggested an approach to classifying multi-label field text in a different field. A semi-automatic creation of the presence structure of existence is applied to uncover the most prominent concepts in the field to which the documents refer, and thus these concepts are used as potential topics for texts whose content is not known in advance. A multiple label task is converted to one or more individual naming tasks. A conversion policy can be used with an arbitrary workbook that results in a probability distribution of categories. This was carried out using vector support machines, the resolution tree, and Naïve Bayes, the nearest neighbor to evaluate the transformation methods. Through direct interaction with multi-tag data, the correlation between different labels was achieved. The multi-label neural network and the "*k*-nearest neighbor algorithms" were utilized.

Q. He et al. [22] provided a framework for classifying text with an existential support for structure and classification of text documents. The ontology structure provides the extraction of structured elements of text that provides them with vocabulary knowledge of relevant information. Structural Ontology helps produce well-organized output documents that represent a specific vector region. In order to avoid the ambiguity usually encountered in semantics, ontological support is provided as logical rules. Naïve Bayes and Ripper algorithms are applied to automated learning in an architecture to build the classifier and classify unknown text documents.

Classification techniques [26], [27] showed promising performance in the development of existing mining research on the web. Web data service information in the form of data returns the document that is classified at run time with its links, and other elements of web pages that are difficult to understand and process due to limited knowledge. In this aspect, knowing the domain service in categories or the typical semantics class will be able to resolve domain service classification restrictions. In this research, it is developed by designing a "hierarchical model of related terms" for domain services with regular updates automatically.

III. PROPOSED SEMANTIC BASED TERMS RELATION APPROACH

Most online services that promote their information contain a large amount of information on the web and are described in natural language, so it will be unclear to address the exact relevance of different users' queries. Additionally, the online service information does not contain a specific format or criteria, and it differs from another Web Services page. To make these pages effectively relevant and accurately categorized, this paper propose a Semantic based Terms Relation Approach (STRA) for classification.

The proposed STRA for web information of services classification works on two phases as shown in Figure-1. The STRA consist of two basic functionality of the classifiers as, Learning and Classification. To do so, it initially performs the "Related Terms Hierarchal Model (RTHM)" construction to learn the relation among the service terms utilizing a "Concept Terms Similarity (CTS)" method. Later, through a modified "Naive Bayes" mechanism [7] it implements a classification methodology.

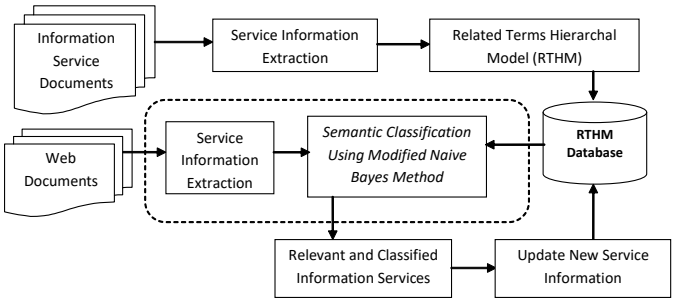

Figure-1: STRA framework for Web information of services classification

The learning knowledge of relevant terms provides efficient knowledge of the classification of diverse and disparate service information that is implemented in a fundamental manner through CTS. The RTHM database will support the efficient classification of relevant documents.

3.1 Learning Methodology and Construction of RTHM

A. Data Extraction

Data extraction from the web pages is the primary process. It parse the each individual documents to extract the defined metadata information in the page. The extracted metadata are tokenized into terms sets. Each set of terms are process to filtered out the stop or junk words and special characters.

The final extracted terms are utilized to construct the required RTHM for further need of classification.

B. Construction of RTHM

To construct an organized RTHM of service it should have some existing domain service knowledge. Each web service page is related to a class of service known as "*Subject*", and the terms which sub-categorized a subject into further small domain categories known as "*Sub-subject*", and the terms which describing the sub-subjects known as "*Topics*". Utilizing the knowledge of the existing domain services it relates the extracted terms and construct a relevance terms hierarchal structure as shown for a service in Fig.1.

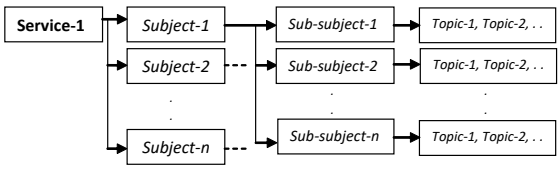

Fig.2: Relevance Terms Hierarchal Structure

The RTHM are the references for the more than one service information which might be related in some manner. Such kind of relevance hierarchical domain knowledge are being used in the past for web data classification [16]. As, X. Zhong et al. [20] proposed a domain hierarchical model for mining in terms of understanding the domain concept and continuous learning schemes of the web pages. Even the Navigli et al. [21] describes the "OntoLearn approach" for understanding the concepts and relationships among the documents, and Jiang and Tan [5] presents a content mining method through learning various domain-specific documents.

The information collected for each service in the form subjects and terms are utilized to construct the RTHM. The construction of the RTHM initiated through identifies *t* kinds of "Top Level Terms (TLT)" of the model. The TLT are the subjects which are considered as primary terms of the services. Each TLT consists of *n* number of sub-subjects categories identifies as "Sub-Level Terms (SLT)", and each SLT consists *k* number of terms in relevant identifies as "Topics Terms (TT)" from the extracted domain web pages.

It RTHM is being represented in three level of vector array. The primary TLT vector is represented as, "*V_TLT*", the SLT vector is represented as, "*V_SLT*" and the vector of TT is presented as, "*V_TT*". Each terms in the vector level is presented as given in the Table-1.

Table-1: Terms Vector Representation for RTHM

Term Level	Vector Terms Representation
V_TLT	*{ ('p','SLT_Term_1'), ('p+1',' SLT_Term_2') , ... , ('p+t','SLT_Term_n') }*
V_SLT	*{ ('p', 'q', 'TT_Term1_1'), ('p+1', 'q+1', ' TT_Term_2') , ... , ('p+t', 'q+n', ' TT_Term_k') }*
V_TT	*{ ('p','q','r','Term_1'), ('p+1','q+1','r+1',' Term_2') , ... , ('p+t', 'q+n', 'r+k', 'Term_k') }.*

This term levels can able to grow at any level as per the new service information addition in the web automatically. These RTHM categorization will be used to classify the online web services through a Concept Terms Similarity (CTS) method as discussed in next section.

3.2 Concept Terms Similarity (CTS) Method

The CTS method aims to associate the documents through calculating the semantic relevance [6], [7] to the RTHM term level categorization. To relate the concept in a document to the subject or primary term level "*V_TLT*" of RTHM, it is needed to process the metadata of the document. So, to perform a document association, it implements a mechanism similar to the defined "Concept Similarity mechanism" [28]. The concept of these words, phrases or metadata finds its relevancy. It is also related to the other elements of the domain to which it belongs.

Let's define *W* is a set of domain pages, where W_i is each individual domain page. Each W_i consists of *n* number of terms represented as, "$W_i = \{t_1, t_2, \ldots, t_n\}$". In order to indentifying the relation between the domain metadata and W_i the CTS compute the frequent terms calculation as, *freq(t)* using Eq. (1) and the term relevancy association in related to constructed RTHM levels.

$$freq(t) = \frac{T_n}{G}, T \in S_k. \quad (1)$$

In the Eq. (1), the "*G*" define the complete terms extracted from the document of WIS, "*T* " define the relevant terms present in the comparing document, and "S_k" define the relevant terms present in RTHM in domain relevancy. Later, to find the probability of similarity relevance we compute the "*P(t|s)*" , where *t* is the set of terms in the domain documents, and *s* is the set of present in RTHM using the Eq. (2).

$$sim(d:s) = \frac{Prob\ (t \cap s)}{Prob\ (s)} \quad (2)$$

For example, if in a document we have a phrase likes "*Electronic Sales*". Base on above association we do the following to finding the concept similarities,

- Compute $\frac{P(Electronic \cap Digital\ Marketing)}{P(E-Commerce)}$ i.e. *P(Electronic | Digital Marketing)*
- Compute $\frac{P(Sales \cap Digital\ Marketing)}{P(E-Commerce)}$ i.e. *sim(Electronic : Digital Marketing)*

The computed "*sim*" values are passed to *T x S formation* as shown in Table-1 below.

Table-1: *T x S* similarity computation

	s_1	s_2	...	s_n	$\sum$ values
t_1	$sim(t_1, s_1)$	$sim(t_1, s_2)$	...	$sim(t_1, s_k)$	$\sum_{n=1}^{T} sim(t_1, s_k)$
t_2	$sim(t_2, s_1)$	$sim(t_2, s_2)$	...	$sim(t_1, s_k)$	$\sum_{n=1}^{T} sim(t_2, s_k)$
...	...	...	...	...	...
t_n	$sim(t_n, s_1)$	$sim(t_n, s_2)$	...	$sim(t_n, s_k)$	$\sum_{n=1}^{T} sim(t_n, s_k)$

Here, T is the number of terms extracted and S is the number of RTHM data. The sum of values are the order more based on sorted to select and the results are considered classified document against the RTHM term level categorization.

3.3 Semantic Classification using RTHM

The process of semantic classification implements the method of CTS using RTHM. The CTS method derives the mechanism of "Naïve Bayesian Classifier (NBC)" [9] to recognize the probable semantic relation. The CTS method initially as assumes that the extracted features terms are independent from a class due to it variance in values, which is defined as condition independency in NBC to make simplification in individual relation computation. It makes to relate the class in support of the computation relation dependency among the features terms. The NBC mechanism is effective if the information system provide accurate supportive feature to provide accurate results, but this dependency and the probability hypothesis can result lot of incorrect classification if small variation in the input information. To overcome this limitation the CTS based semantic classification with the variable relevant terms as RTHM to efficiently classifies. The Algorithm-1 present the steps of computation flow.

Algorithm-1: CTS based Semantic Classification

Input:
T [] → Set of Web information service Terms
Information_RTHMSet[ ][ ] → *n-dimensional vector of Information RTHM[ ]*
Information_Terms[ ] → Set of Service Terms from *Information_RTHMSet[ ][ ]*

For each term t_i of *T[]*
  For of each *Information_RTHMSet* $[c_i][\,]$
    Information_Terms[ ] ← *Information_RTHMSet[x][y]*
      For each Information term of s_i in *Information_Terms [k]*
        Compute the CTS probability resemblance, β of s_i in *Information_Terms [k]*
      For End
  For End
  If $\beta >= 1$ **then**
    t_i, *relevant class* → c_i
  If End
For End

If the semantic classified web information service have Bayes probability similarity, $\beta >= 1$ then we can decide the classified service are highly related. To evaluate the above proposal we performs a extensive analysis over more than 500 different WIS pages collected from different domain. The STRA integration mechanism will be improvising the accuracy of classification using the define services in the RTHM.

IV. EXPERIMENT EVALUATION & RESULTS

4.1 Datasets

We are using the WIS metadata to generate a synthetic data for the experimental investigation of the RTHM structure as shown in Figure-3 and also collected 500 different WIS pages from different domains for test database. The domain information are collected from well known service related to four domains as "Online Sales", "Healthcare", "Online Booking" and "Entrainments" from Open Directory Project (http://www.dmoz.org/). The constructed RTHM model will support in service classification.

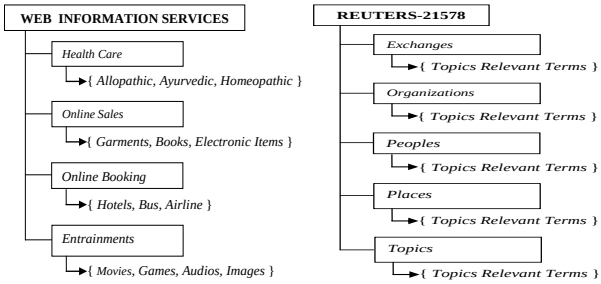

Figure-3 : RTHM Model of WIS and Reuters-21578 Datasets

To extend the evaluation in a dataset collection we choose a Reuters-21578 Datasets. It consists of 22 SGML-DTD data file and each file consisting of 1000 records, and each records describing data format related to six categories which is used to index the data for text categorization. The datasets provide a classification task with challenging properties having multiple categories with overlapping and non exhaustive relationships among the categories [28]. To evaluate and measure the system we implement a series of experiments to evaluate the STRA with synthetic and *Reuters-21578 Datasets* and perform the Precision and Recall calculations for the evaluation.

4.3 Result Analysis

A. Synthetic Data Results

The analysis made using synthetic data using different domain queries with a variation of association threshold value from 10 to 100. The related classified web information service list shown in Table-2.

Table-2: Classified WIS in different Services

Top Level Terms (TLT)	Sub-Level Terms (SLT)	Classified Web Information Services
Entrainments	Movies	https://erosnow.com/movies/
Entrainments	Games	https://play.google.com/store/
Entrainments	Audio	https://mp3download.center/
Online Sales	Garments	https://www.amazon.in/Clothing-accesories/
Online Sales	Books	https://www.indiabookstore.net/
Online Sales	Electronic Items	https://www.flipkart.com/
Healthcare	Allopathic	https://www.verywellhealth.com/
Healthcare	Ayurvedic	https://www.ayurvedichealing.net/
Healthcare	Homeopathic	https://www.drbatras.com/homeopathy/
Online Booking	Airline	https://www.makemytrip.com/flights/
Online Booking	Bus	http://www.redbus.in/
Online Booking	Hotel	https://www.trivago.in/

Figure-4 and 5 describes the precision and recall performance against the relevancy threshold of STRA. The RTHM based classification shows an improvisation precision with low recall against different relevancy threshold. Relevancy threshold is a value which controls the filter of relevance document for a query. The higher the threshold the higher the precision, which proves the improvisation of the proposal.

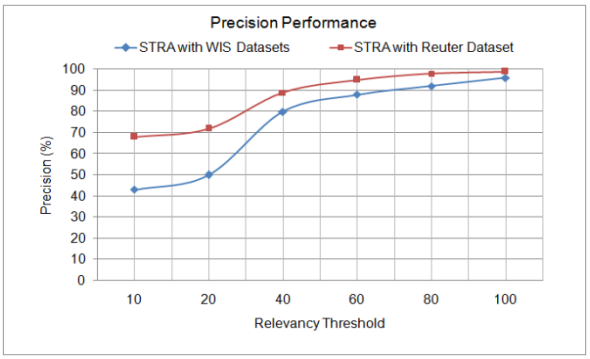

Figure-4: Precision Vs Relevancy Threshold

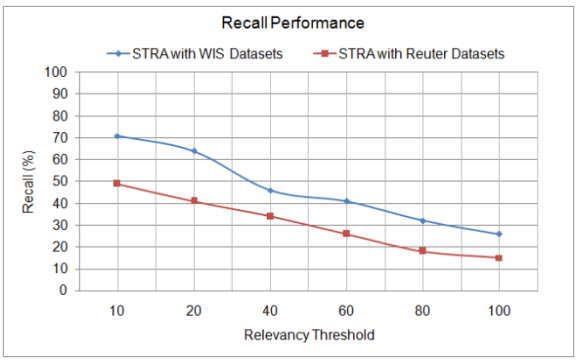

Figure-5: Recall Vs Relevancy Threshold

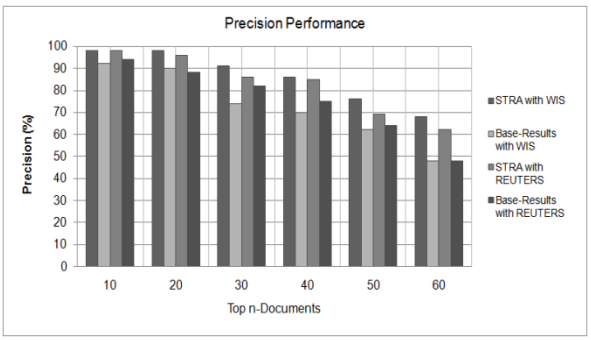

Figure-6: Precision Vs Top n-documents

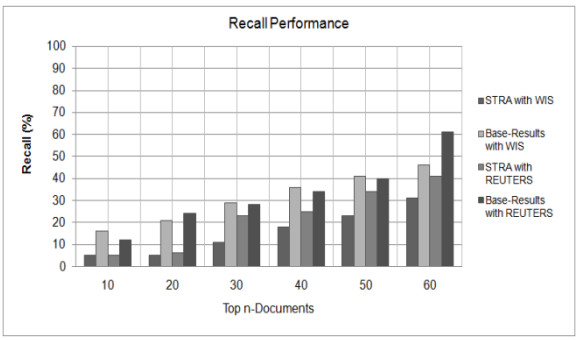

Figure-7: Recall Vs Top n-documents

Figure-6 and 7 describes the precision and recall measures of STRA in compare with base search result. It shows result in compare with top n-documents retrieved for a query. It shows a better precision against base search result due to the support of RTHM knowledge and also runtime updating also improvise it further.

B. Reuters-21578 Datasets Results

To evaluate with the extensive performance of STRA we compare the classification with existing "ADTree" and "NBTree" classifiers using *Reuters-21578* Datasets. The experiment evaluation measure "True positive", "False positive", "True negative" and "False negative" to compute the performance of the classifiers as shown in Figure-8. It shows an average of 25% of improvisation in compare to the existing classifiers.

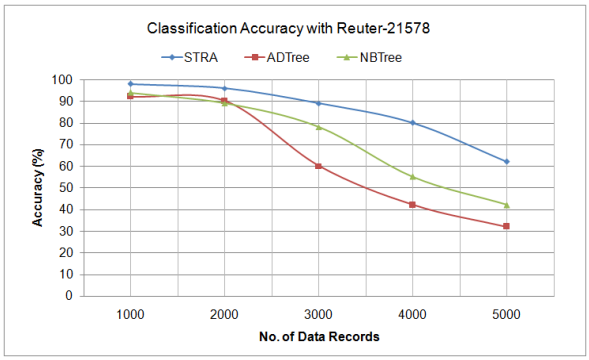

Figure-8: Classification Accuracy Comparison *Vs Reuters* Dataset Records

V. CONCLUSION

This paper presents a Semantic based Terms Relation Approach in a Web environment for classification of Web information service using RTHM support. Integration of semantic classification using the Concept Terms Similarity and the Naïve Bayes classification. The semantic classification calculates the relevance association to relate the semantic relevancy. It performs the document correlation, through the Concept Terms Similarity and classification through Naïve Bayes classification approach to ensure the accuracy classification of the results. This approach is similar to the semantic association function where the relationship between the RTHM term levels data and the service metadata. The experiment results show an increase in the value of the precision, and reduction in the irrelevant information in compare to the existing classifiers. A more comparative analysis of the works will be performed in our feature works.

REFERENCES

1. F. Naderkhah, H. Moradi, "A Website Analytics System Considering Users' Category", IEEE 5th Conference on Knowledge Based Engineering and Innovation (KBEI), 2019.
2. J. Shen, E. Zheng, Z. Cheng, C. Deng, "Assisting Attraction Classification by Harvesting Web Data", IEEE Access Volume: 5 Pages: 1600 - 1608, 2017.
3. Z. Wang, B. Song, "Research on hot news classification algorithm based on deep learning", IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2019.
4. F.-Dombeu, J., Huisman, M., "Combining ontology development methodologies and semantic web platforms for e-government domain ontology development", International Journal of Web & Semantic Technology (IJWesT), Vol.2(2), pp.12-25, 2011.
5. A. Sieg, B. Mobasher, Robin Burke, "Web Search Personalization with Ontological User Profiles", ACM CIKM'07,isboa, Portugal, November 6-8, 2007.
6. T. -Y. Chan, Y.-S. Chang, "Enhancing Classification Effectiveness of Chinese News Based on Term Frequency", IEEE 7th International Symposium on Cloud and Service Computing (SC2), Pages: 124 - 131,2017.

7. J. Ruohonen, "Classifying Web Exploits with Topic Modeling", 28th International Workshop on Database and Expert Systems Applications (DEXA) Pages: 93 - 97, 2017.
8. L. Yao, Quan Z. Sheng, Anne. H.H. Ngu, Jian Yu, Aviv Segev, "Unified Collaborative and Content-Based Web Service Recommendation", IEEE Transactions On Services Computing, Vol. 8, No. 3, May/June 2015.
9. M. Shirakawa, K. Nakayama, T. Hara, Shojiro Nishio, "Wikipedia-Based Semantic Similarity Measurements for Noisy Short Texts Using Extended Naive Bayes", IEEE Emerging Topics in Computing, Volume 3, No. 2, June 2015.
10. S. Vogrincic, Z. Bosnic, "Ontology-based multi-label classification of economic articles", Computer Science and Information Systems, Vol.8(1), pp.101-19, 2011.
11. A. Harth, M. Janik, S. Staab, "Semantic Web architecture", in Handbook of Semantic Web Technologies, J. Domingue, D. Fensel, and J. A. Hendler, Eds. Berlin, Germany: Springer-Verlag, pp. 43-75, 2011.
12. X. Tao, Y. Li, N. Zhong, R. Nayak, "Ontology Mining for Personalized Web Information Gathering", Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence, pp. 351-358, 2007.
13. H. Dong, F. K. Hussain, "Self-Adaptive Semantic Focused Crawler for Mining Services Information Discovery", IEEE Transactions On Industrial Informatics, Vol. 10, No. 2, May 2014.
14. K. Belhajjame, S. M. Embury, N. W. Paton, "Verification of Semantic Web Service Annotations Using Ontology-Based Partitioning", IEEE Transactions On Services Computing, Vol. 7, No. 3, July-september 2014.
15. A. Rozeva, "Classification of text documents supervised by domain ontologies", Applied Technologies & Innovations Volume-8, Issue-3, pp.1-12, November-2012.
16. R. Mohanty, V. Ravi, M. Patra, "Classification of Web Services Using Bayesian Network", In Journal of Software Engineering and Applications, pp. 291-296, 2012.
17. M. Janik, K. Kochut, "Wikipedia in action: Ontological knowledge in text categorization", IEEE International Conference on Semantic Computing, pp.268-75, 2008.
18. E. Al-Masri, Q.H. Mahmoud, "Investigating Web Services on the World Wide Web", Proc. 17th Int'l Conf. World Wide Web (WWW '08), Apr. 2008.
19. T. Tran, P. Cimiano, S. Rudolph, R. Studer, "Ontology-Based Interpretation of Keywords for Semantic Search", Proc. Sixth Int'l Semantic Web and Second Asian Semantic Web Conf. (ISWC '07/ ASWC '07), pp. 523-536, 2007.
20. X. Jiang, A.H. Tan, "Mining Ontological Knowledge from Domain-Specific Text Documents", Proc. Fifth IEEE Int'l Conf. Data Mining (ICDM '05), pp. 665-668, 2005.
21. A. Singh, K. Nakata, "Hierarchical classification of web search results using personalized ontologies", In Proceedings of the 3rd International Conference on Universal Access in Human-Computer Interaction, HCI International 2005, Las Vegas, NV, July 2005.
22. Q. He, L. Qiu, G. Zhao, S. Wang, "Text categorization based on domain ontology", Web Information Systems - WISE 2004, Vol.3306, Springer Verlag Berlin Heidelberg, Germany, pp.319-24, 2004.
23. N. Zhong, "Representation and Construction of Ontologies for Web Intelligence", Int'l J. Foundation of Computer Science, vol. 13, no. 4, pp. 555-570, 2002.
24. R. Navigli, P. Velardi, A. Gangemi, "Ontology Learning and Its Application to Automated Terminology Translation", IEEE Intelligent Systems, vol. 18, no. 1, pp. 22-31, Jan./Feb. 2003.
25. Z. Duo, L. Zi, X. Bin, "Web service annotation using ontology mapping", In IEEE International Workshop on Service-Oriented System Engineering, pp. 235-242, 2005.
26. M. Corella, P. Castells, "Semi-automatic semantic-based Web service classification", In Business Process Management Workshops, volume 4103 of LNCS, Vienna, Austria, Springer, pp. 459-470, 2006.
27. Y. Li, N. Zhong, "Mining Ontology for Automatically Acquiring Web User Information Needs", IEEE Trans. Knowledge and Data Eng., vol. 18, no. 4, pp. 554-568, Apr. 2006.
28. S. Sekine, H. Suzuki, "Acquiring Ontological Knowledge from Query Logs", Proc. 16th Int'l Conf. World Wide Web (WWW '07), pp. 1223-1224, 2007.
29. M. Bruno, G. Canfora, M.D. Penta, R. Scognamiglio, "An Approach to Support Web Service Classification and Annotation", Proc. IEEE Int'l Conf. E-Technology, E-Commerce and E-Service (EEE '05), 2005.
30. H. L. Johnson, K. B. Cohen, W. A. Baumgartner, Z. Lu, M. Bada, T. Kester, H. Kim, L. Hunter, "Evaluation of Lexical Methods for Detecting Relationships Between Concepts from Multiple Ontologies", Proc. Pacific Symp. Biocomputing, 2006.
31. N. Oldham, C. Thomas, A. Sheth, K. Verma, "METEOR-S Web Service Annotation Framework with Machine Learning Classification", Semantic Web Services and Web Process Composition, vol. 3387, pp. 137-146, Jan. 2005.
32. A. Doan, J. Madhavan, P. Domingos, A. Halevy, "Learning to Map between Ontologies on the Semantic Web", Proc. 11th Int'l Conf. World Wide Web (WWW '02), pp. 662-673, 2002.