

Ensemble based Machine Learning using Ontology Information Extraction for Information Retrieval

Ritesh Kumar Shah, Sarvottam Dixit



Abstract: *The movement of internet technology enhanced the speed and accuracy of data retrieval over the internet. The retrieval of data over the internet needs some automatic process of information extraction and query retrieval. The information extraction gives the process of the predefined structure of the concept to a particular domain of knowledge. The process of information extraction proceeds in two steps one is pre-processing of data and post-processing of data. In preprocessing of data used the concept of the glowworm optimization algorithm. The glowworm algorithm is a family of kits a gives the better selection of information in constraints of similarity. The selection of similarity based on the process of lubrication. The optimization of glowworm removed the unwanted noise of data and filtered it. For the extraction of information used ensemble-based information extraction. The ensemble-based information extraction proceeds with constraints function that function is called mapper constraints. The mapper constraints map the process of ontology with guided domain ontology. The ensemble-based information extraction process used the concept of machine learning for the binding of process. The goals of this work are the development of an OBIE for the domain of different fields of data retrieval such as news agencies, hotel industries and sports. The proposed model combines with the use of ontology, POS and language processing tools and constraints-based mapper with domain ontology.*

Keywords: *Web Mining, Semantic Web Mining, Ontology, Information Extraction, SVM.*

I. INTRODUCTION

Semantic Web Mining represents combining the two fast-developing research areas Semantic Web and Web Mining. The main aims of Web Mining to extracting the relevant information, knowledge discover from unstructured web data [1][3]. With the rapid development and wide application of the internet, Web has a huge amount and complex information on the World Wide Web. Nature of these documents concerned; they are not descriptive enough to express themselves, excess and distributed information all over the World. So, it becomes too complicated and retrieval for all the needed data from the web. Semantic web [2][10] is the solution to this problem. The idea behind the semantic Web is to make the Web as intelligent as possible. The success story behind the semantic web is ontology. Ontology creates the relationship between concept and relation.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Ritesh Kumar Shah*, Research Scholar (CSE), Mewar University, Gangarar, Rajasthan, India.

Prof (Dr.) Sarvottam Dixit Research Guide, Mewar University, Gangarar, Rajasthan, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

The formal representation allows the Web Documents to be more understandable for the machine as well as an individual [4]. Ontology-based information extraction process derived by the information extraction process of semantic web-based information. The extraction of information based on ontology proceed in three directions first one is natural language processing, the second one is the process of statistical methods such as entropy-based and probability-based methods and finally used machine learning algorithms for the extraction of information based on ontology[7][9]. The machine learning-based algorithms used in information extraction in limited scopes such as support vector machine and decision tree. These methods of word correction are low. In the derived concept of machine learning, another method is called ensemble classifier. The ensemble classifier-based information extraction improved and enhanced the process of information extraction. The glowworm algorithms work in the principle of luciferin. The luciferin collects similar data and discards dissimilar data (the process of noise elimination). The similar data pass through the process of parts of speech (POS). The POS process creates the segmentation of tokens and keywords and generation feature vector for the processing of ensemble-based information extractor[1] [19]. The ensemble-based information extractor four grouped categories such as concept, properties, the relationship of class and value (C P R V). the value of C and P classified with ensemble classifier and the value of R and V generates through the constraints defined by the ensemble. The rest paper, in this section-II described literature survey, in the section-III, I discussed information extraction, in the section-IV implementation and result discussion and finally discuss conclusion and future work in section-V.

II. LITERATURE SURVEY

Farman Ali, Daehan Kwak, Et al. [1] In this paper, they talked about a combined ontology and SVM-based suggestion and data extraction framework. The framework robotizes the extraction of exact information from the Internet and proposes precise things for incapacitated clients. The examined framework improves the exhibition in suggestions and data extraction. Their examined framework can successfully help impaired clients to recover data from website pages effortlessly, classifies the content. Afterward, recommends the right things, since it can dispense with clamor, remove the thing highlight data and assumption words, and sort the notion words into an all-inclusive level of extremity terms.

Farman Ali Et al. [2] this paper talked about a powerful classification strategy for highlight survey's identification and semantic information for opinion mining dependent on SVM and Fuzzy Domain Ontology (FDO). The talked about framework recovers an assortment of audits about lodging and inn highlights. The SVM identifies include surveys and filter out superfluous audits and the FDO is then used to register the extremity term of each element. The amalgamation of FDO and SVM significantly builds the exactness pace of audit and assessment word's extraction and precision of opinion mining. The FDO and keen model are created utilizing Protégé OWL-2 device and JAVA, individually. The experimental result shows impressive execution improvement in highlight the audit's classification and opinion mining.

Zhen-Shu Mi Et al. [3] In this examination, they presented a hybrid mechanism dependent on fuzzy domain ontology and SVM to accelerate submerged obstruction acknowledgment. A few scientists have tended to this issue and a considerable lot of them discussed presented answers to take care of this issue. To examine the presentation of their talked-about framework, they built up a model framework and played out various investigations by utilizing the preparation set from the COIL database. To screen the general execution of the framework, they contrasted their examined plan and BP and exemplary SVM plans. Their examined plan demonstrated increasingly precise outcomes in a shorter range of time.

O. Abuomar, S. Nouranian Et al. [4] The planned SVMs model is resourceful for materials researchers and specialists. Since it tends to be utilized to subjectively survey diverse nanocomposite mechanical reactions related to various blends of the definition, handling, and ecological conditions. All in one, the SVMs model utilizing the re-substitution technique had the option to more readily anticipating the ideal property classes with an insignificant AE rate utilizing the dab item and degree two polynomial bits than the 3-overlay CV strategy. Above all, the created SVMs model can recognize the mechanical property reaction esteem coming about because of a chose untested mix of the nine input factors referenced in this investigation. The decision of the info level mixes is comparable with specific ideal measure(s) considered by materials researchers and architects.

Maciej Zieba Et al. [5] They have talked about a novel SVM technique for imbalanced information issue, which was additionally utilized for rules extraction. They have assessed the nature of the talked about the approach by contrasting it and different arrangements committed for imbalanced information issues. Next, they have utilized the talked-about technique to take care of the issue for the expectation of the post-employable future in lung malignancy patients. They have indicated that their methodology can be effectively applied to the issue by making extra test examinations on the genuine dataset. At last, they extricated choice guidelines utilizing prophet-based methodology.

Zehra ,Camlica1, H.R. Tizhoosh and Farzad Khalvati Et al. [6] In this paper, they talked about to identify notable areas of pictures during preparing and overlay the information to diminish the impact of immaterial districts. Accordingly, littler picture regions will be utilized for LBP highlights estimation and therefore, grouping by SVM. they use IRMA 2009 dataset with 14,410 x-ray pictures to

confirm the exhibition of the talked about the approach. The talked about restorative picture grouping utilizing a saliency-based collapsing strategy gives off an impression of being a viable technique when bolster vector machines and nearby parallel examples are utilized. Collapsing non-remarkable pieces of the picture may bring about the slight increment of grouping blunder. That might be normal since collapsing territories cover with notable districts bringing about slight bending. The decision on how to overlay picture squares is the most basic piece of the pre-preparing.

Ignas Kukenys and Brendan McCane Et al. [7] They construct cascades of such approximations and use them for object discovery, estimating their capacity to reject non-objects and their normal time taken early. They, at that point, propose a half and half approach, which joins the two strategies and further improves the exhibition of the SVM course. They present recommend that a universally enhanced decreased set as proposed by Burges offers higher precision while approximating a single SVM work, anyway in the course situation, the ravenous methodology offers a huge increment in speed because of the sharing of the support vectors between capacities in the course.

Esraa Elhariri Et al. [8] The examined methodology comprises of three stages; specifically, pre-handling, include extraction, and classification stages. Since tomato surface shading is the most significant trademark to watch readiness, this framework utilizes a hued histogram for arranging readiness organize. They built up a framework for characterizing the readiness phases of tomato. The framework has three fundamental stages; pre-preparing, include extraction and readiness classification. The work was practiced by applying resizing, foundation evacuation, and separating shading parts to each picture at that point include extraction is applied to each pre-prepared picture, HSV histogram and shading minutes are gotten as a component vector and utilized as a PCA contribution for change. fiFinally SVM model is created for readiness to organize classification.

Wenjia Li Et al. [9] In this paper, they study a malware detection plot for the Android stage utilizing an SVM-based methodology, which coordinates both hazardous consent blends and helpless API calls and uses them as highlights in the SVM calculation. To approve exhibition of examined methodology, broad tests have been directed, which show that the talked about malware recognition plot can recognize malignant Android applications successfully and proficiently. They talked about an SVM-based malware location plot for the Android stage and utilized both hazardous API calls and unsafe authorization mixes as highlights to assemble an SVM classifier, which can consequently recognize vindictive Android applications (malware) from authentic ones. Analysis results show that the talked about the plan can identify malware in a precise way. Wei Zhou Et al. [10] study the utilization of SVM based strategy and gathering qualities in assault profiles. A two-stage recognizing strategy SVM-TIA is examined dependent on these two strategies. They talked about a strategy to distinguish peddling assaults dependent on SVM and target thing investigation technique — trials on the Movie Lens Dataset exhibit the adequacy of SVM-TIA in distinguishing pushing assaults.

The constraint for SVM based recognition strategy is that class unbalances issue exists in SVM classifier. Marginal SMOTE strategy is utilized to reduce the class unbalance issue. An SVM classifier joined with the Borderline-SMOTE technique utilizing profile extraction qualities is examined. On the opposite side, the target thing investigation technique is utilized to decrease the false-positive rate of the recognition result.

Harsh Arora and Govind Murari Upadhyay Et al. [11] presented the existing concepts, for the identification of suspicious activities. The considered author's work concentrates on the use of concepts like CrossSpot Algorithm, The Twitter Grapevine: Web Application, Traditional Corpus-based approach, Similarity Distance-based Differentiation Approach, KNN approach in Crime Detection and Criminal Identification Concept, Crime Data Mining, Suspicious Pattern Detection Algorithm, Latent Sentiment Analysis, etc. Due to various research gaps in the existing concepts, they have presented an integrated approach of SVM and PSO algorithm for the detection of suspicious activities on online forums. The discussed framework presents the workflow, as mentioned.

Agnieszka Konys Et al. [12] The paper is a fruitful endeavor to the state-of-the-art of chose OBIE frameworks, trailed by the procedure of scientific categorization development and consequence information systematization of specific OBIE approaches. In fact, ontology gives information that can be joined into any database, information base, or data framework holding information related to OBI. The formalized structure of the cosmology offers machine-readable access and taking care of semantic information is an intriguing advance to improve the looking through limit and information sharing of the talked about philosophy. This type of critical solving guarantees semantic interoperability for information and information gathered.

Kamran Munir and M. Sheraz Anjum Et al. [13] This paper examine ontology-based data recovery approaches and strategies by taking into consideration the parts of philosophy displaying, preparing and the interpretation of ontological information into database search demands. This examination helps in deciding the social elements that might be changed into philosophy ideas. It likewise bunches together or separates in events the data indicated in a social database table and to decide connections between various tables. The greater part of the current database-to-ontology change approaches doesn't give an accurate portrayal of the space metadata in cosmology and doesn't empower the age of the separate database relations.

Diana Maynard, Milena Yankova, Alexandros Kourakis, and Antonis Kokossis Et al. [14] In this paper, they have displayed an application for programmed information extraction, the board and checking in the Chemical Engineering area, incorporated in unique information the executive's entrance. Joined with different instruments and applications for information building found inside the entrance, it shapes the premise of a framework for data recovery, wording securing and innovation watch. The application has been tried in the Employment division with phenomenal outcomes and has been effectively ported to different classifications of content, for example, news things and friends reports.

Jie Chen, Chunxia Zhang and Zhendong Niu Et al. [15] talk about a two-advance resume data extraction approach. In the initial step, the crude content of the resume

is recognized as various resume squares. To accomplish the objective, they structure a novel component, Writing Style, to display sentence language structure data. Other than word file and accentuation file, lexical word trait and expectation consequences of classifiers are remembered for Writing Style. At that point, various classifiers are utilized to distinguish various qualities of actuality data in resumes. Test results on a genuine world dataset show that the calculation is doable and compelling. They talked about a Writing Style to recognize various lines. Contrasted with those removing calculations, in light of either HMM or CRF, their methodology doesn't require an excessive amount of physically clarified preparing set, which can spare bunches of human endeavors and time.

Juan F. Sequeda Et al. [16] An effectively rehashed use case for Semantic Web advancements in Ontology-Based Data Access for information incorporation. Here, two challenges for building up an OBDA framework are considered: philosophy and mapping designing, alongside a pay-more only as costs arise strategy that tends to these difficulties and empowers readiness. This pay-more only as costs arise technique is an initial phase in demonstrating how ontologies and mappings can be designed in a deft way for OBDA. As far as they could possibly know, the building of metaphysics and mappings for OBDA is as yet an open territory for inquiring about. Given that R2RML mappings are spoken to in Resource Description Framework (RDF), these can be put away in a triple-store, questioned, and contemplated upon.

Kamel Nebhi Et al. [17] a standard-based framework for the acknowledgment and semantic disambiguation of named elements in tweets. As their exploratory outcomes appear, execution of this methodology estimated through BDM looks encouraging when utilizing Linked Data as Freebase and linguistic setting for disambiguation. They presented a methodology for Ontology-based Information Extraction from Twitter. Their framework gives an incorporated disambiguation module dependent on fame score and punctuation-based similitude. As their assessment appears, the framework performed better-utilizing disambiguation process altogether.

Xiaomeng Su and Jon Atle Gulla Et al. [18] The examined usefulness encourages the mapping procedure in that the semantic heterogeneity and the hazard for peculiarities are diminished. Moreover, the psychological over-burden of the client is diminished in the manual piece of the mapping procedure. The methodology can be applied in a few other various settings, in view of space free nature of the methodology. One such setting is archives recovery and distribution between various online interfaces. Clients may comply with their nearby ontologies through which the online interfaces are sorted out. It is attractive to have support for robotized trade of records between the gateways and still let the clients keep their viewpoints. Administration coordinating is one more possibility to apply the technique, however, they should expect that there are some assistance portrayal progressive systems and that the supplier and the requester are utilizing diverse arrangement constructions.

Daya C. Wimalasuriya Et al. [19] This paper displays the subtleties of some execution work completed by the creator to investigate the utilization of philosophy-based data extraction.

These investigations were of two kinds - evaluating ebb and flow look into chip away at the field and doing a few executions to investigate the field. Checking on the flow inquire about work empowered me to detail a definition for an OBIE framework, to distinguish a typical design for OBIE frameworks and to arrange existing OBIE frameworks along with their key measurements. The usage work enabled me to utilize the information got on OBIE frameworks by investigating ebb and flowchart into work. It additionally enabled me to get acquainted with the generally utilized programming devices and systems of the field. Also, these usages delivered some fascinating outcomes.

Apichai Suesatsakulchai, Marut Buranarach and Orasa Tetiwat Et al. [20] This paper describe a BDCSRS prototype development using an OAM framework to retrieve the complication information of blood donors. The BDCSRS architecture uses a complication ontology that included nine main classes mapped to information that was stored in the database. One hundred cases of blood donor complication information were simulated and used to evaluate semantic information retrieval from the BDCSRS prototype with precision and recall values. The result showed the precision value equal to 1.0 and recall value equal to 1.0. This evaluation result showed that the BDCSRS prototype retrieved the complication information of blood donors accurately. Blood bank staff can, therefore, report complication information to the Nation Blood Centre using information retrieval from the BDCSRS. This information was also beneficial to blood bank staff to reduce the incidence of complications.

Rizwan Iqbal, Masrah Azrifah Azmi Murad, Aida Mustapha, Aida Mustapha, Syed Rizwan Ali and Clay Palmeira da Silva Et al. [21] The utilization of graphical information shown formalisms with a vocabulary agreement understanding of terms of conceptualization of the universe of circle their se is another high potential methodology in the metaphysics designing and information the executives setting. At first, idea maps were utilized in the field of instruction and adapting; however, soon, it got well known in different territories due to its adaptable and instinctive nature. In the field of ontologies, idea maps have been investigated and utilized for encouraging various parts of cosmology advancement. A significant purpose for this inspiration is the auxiliary similarity of idea maps with the various leveled structure of ontologies. This exploration plans to exhibit a quantitative assessment of 2 unique speculations related to the adequacy of utilizing idea maps for metaphysics conceptualization. Results show that idea maps impart the component of learning in the conceptualization procedure.

III. INFORMATION EXTRACTION

The extraction of information based on ontology proceed in three directions first one is natural language processing, the second one is the process of statistical methods such as entropy-based and probability-based methods and finally used machine learning algorithms for the extraction of information based on ontology. The machine learning-based algorithms used in information extraction in limited scopes such as support vector machine and decision tree. These methods of word correction are low. In the derived concept of machine learning, another method is called ensemble classifier[4-6]. The ensemble

classifier-based information extraction improved and enhanced the process of information extraction. The process of information extraction divides into two phases.

1. The processing of input data
2. Generation of knowledge to map with domain ontology

The processing of input data

The processing of input data used a glowworm optimization algorithm. The glowworm algorithms work in the principle of luciferin. The luciferin collects similar data and discards dissimilar data (the process of noise elimination). The similar data pass through the process of parts of speech (POS). The POS process creates the segmentation of tokens and keywords and generation feature vector for the processing of ensemble-based information extractor. The ensemble-based information extractor four grouped categories such as concept, properties, the relationship of class and value (C P R V). the value of C and P classified with ensemble classifier and the value of R and V generates through the constraints defined by the ensemble. The process of ensemble describes here.

The ensemble-based information extractor used the basic property of ontology of categorization. The parameters of categorization describe as concept, property, relationship of class and value. The represents of set is {C, P, R, V}. The processing of ensemble class depends on the feature optimization and feature selection. The process of feature selection used glowworm optimization algorithms. The glowworm optimization creates the divers feature relations of base class and variable class for the process of ensemble classifier.

Consider {Xi} the set of features

EC ensemble classifier

N= number of classifiers

KNN= A classifier

DT= B classifier

SVM =Fc final classifier

GFS= glowworm feature selection

R relationship of data

D= dimension of data

Fs feature space

Algorithm for ensemble

Input: - set of features data {Xi},i=1,2,3,...,M,

$X_i \in R^D$ the dimension of information extraction

Output: - EC {ensemble classifier}

1. Initialize Fs= [R; A, B, Fc] the selection of class depends on GFS
2. For n=1 to N do
3. Estimate Fc={Xi} ∈ (A, B)
4. If f(Xi) == Θ



5. Generate final Ec
6. Else
7. Call GFS
8. End for
9. Return Fc
10. Exit

Modelling of ensemble classifier
 Base class sample data S^D
 Ensemble classifier C_1, C_2, \dots, C_n
 Lema-1(similar feature data)
 $Fc: R^{dA} \rightarrow R^{dB}$ where $D \in R^d$ to set at sample selection of base classifier to ensemble classifier.
 Lema-2 (Dissimilar feature Data)
 $S^D = Fc \rightarrow N(Xi) + (A \cup B)$

Lema-3 (Mutual feature data)

for all classifier $Xi \in S^{di}$ and generates model for class

$$S^D = class\{|S^D - c| + \sum_{i=0}^N |Fc(A) - Fc(B)|$$

Information extraction algorithm IE

Input:- sample data of training
 $S^D = \{s_1, s_2, \dots, s_n\}$ and feature set of realtion $\{xi\}, i = 1, 2, \dots, m$

Output: - IE(Fc)

1 initialize $P=0, IE=\emptyset, FS=\emptyset$

2. start $t=0$

While $i < n$ do

Fc for training data with GFS with predication $Pi = F(S^D), (R^D)$

3. generate new class of information

$$S^D = \{Fc(s^{dA}), (S^{dB})\} \text{ with } S^d \in R^d$$

4. $Pi=i+1$ with S^d

5. estimate $Fci \leftarrow Ai + Bi$

6. count Pi

7. end

8. return $(IE \leftarrow S^D \in R^D)$

Block diagram of methodology

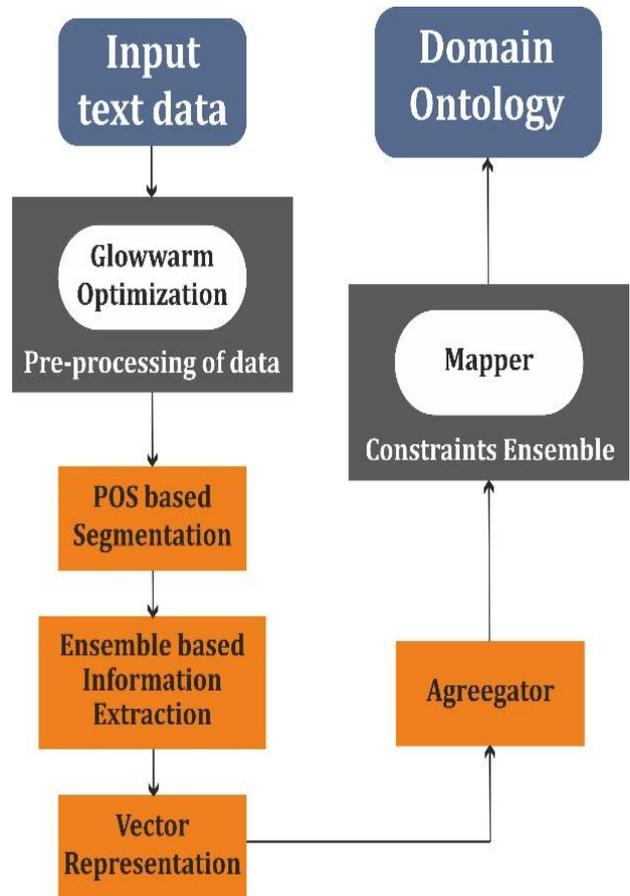


Figure 1: Process model of ontology-based information extraction.

Input text data: - in this module, the data is input for the processing of the query example of Indore is

smart city.

Pre-processing of data: - in this phase of the module, filter the unwanted link and noise data in this phase used glowworm optimization algorithms. The glowworm optimization algorithms used the value of similar and dissimilar and removes noise data.

POS (parts of speech): - in this module used the segmentation of sentences in forms of words. The differentiated words used for the further process of extraction.

EIE (ensemble-based information extraction): - the ensemble-based information extraction is the main module of proposed work. The ensemble information extractor used three machine learning algorithms KNN, DT and SVM. The ensemble algorithms create a predefined class of words based on these algorithms.

Vector Representation: - in this module process, the words for the process of the mapper to map the words with given constraints function. The vector representation input process the classifier for the mapping of domain ontology.

Agregator: - is part of the mapper and vector representation.

Aggregator used the given constraints function for the extracted information by the ensemble information extractor.

Mapper: - in this module, the mapper incorporates the constraints of ensemble information extractor.

Domain ontology: - the domain ontology represents the knowledge information to the respective ontology for the processing of information

IV. SIMULATION & RESULT DISCUSSION

The process of analysis of proposed algorithms used different tools for designing ontology such as protégé OWL and fuski java-based simulator for the creation of the ontology database — the developed knowledge extracted by MATLAB function for ensemble classifier. The evaluation process contains six phases. The phase of evaluation depends on the method of ontology. The validation of the proposed methodology used two standard parameters, such as precision and recall.

Here

TP – true positive,

FP false positive, False negative

$$Precision = TP / (TP + FP) \times 100\%$$

$$Recall = TP / (TP + FN) \times 100\%$$

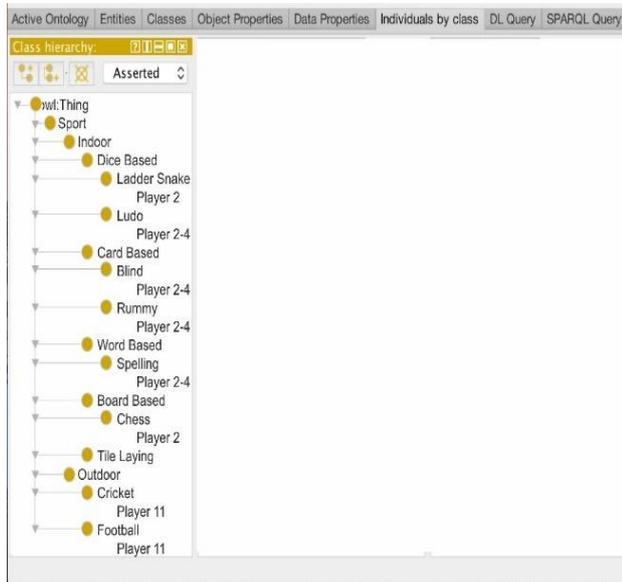


Figure 2: the process of class instance and properties of sport. In the sports ontology, we make a sports root class in owl: Thing. Here we organized the relation of sports class related to the location of sport played like indoor and outdoor. Here discussed its instance of sport name and how many players played these games in this sport ontology.

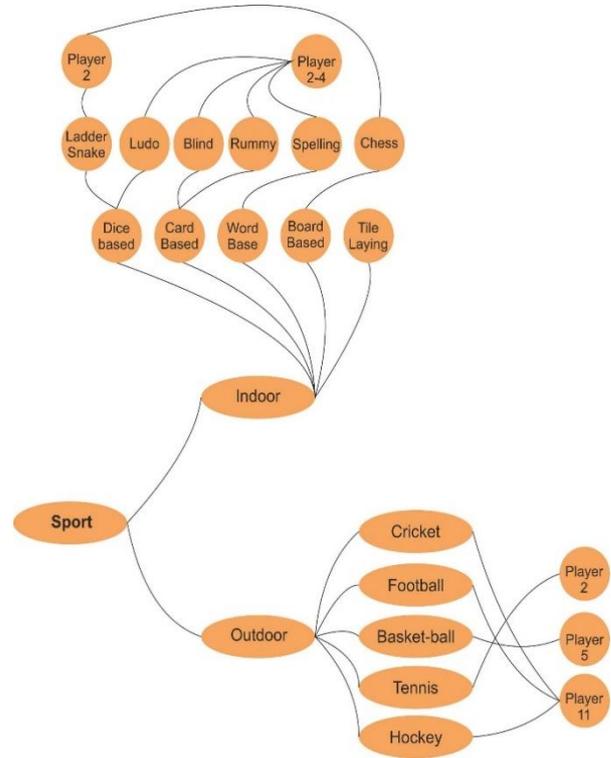


Figure 3: shows the sports ontology generated by Protégé OWL.

| | Precision | Recall |
|---------------------|-----------|--------|
| SVM | 83 | 71 |
| CSVM | 86 | 76 |
| CSVM-Glow | 85 | 82 |
| Used Ontology: News | | |

Table 1: Given the table has shown that the result value of our simulation for News ontology and here used parameter precision and recall shown comparative value but here CSVM-Glow shown better results.

| | Precision | Recall |
|-------------------------|-----------|--------|
| SVM | 68 | 57 |
| CSVM | 76 | 68 |
| CSVM-Glow | 83 | 63 |
| Used Ontology: Research | | |

Table 2: Given the table has shown that the result value of our simulation for research ontology and here used parameter precision and recall shown comparative value but here CSVM-Glow has shown better results.

| | Precision | Recall |
|----------------------|-----------|--------|
| SVM | 69 | 68 |
| CSVM | 75 | 65 |
| CSVM-Glow | 88 | 72 |
| Used Ontology: Sport | | |

Table 3: Given the table has shown that the result value of our simulation for sport ontology and here used parameter precision and recall shown comparative value but here CSVM-Glow has shown better result.

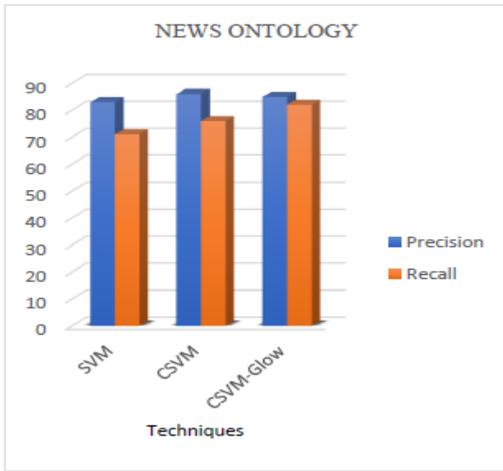


Figure 4: Comparative performance graph showed that the results between SVM, CSVM, and CSVM-Glow techniques for precision and recall parameters in percentage for News ontology.

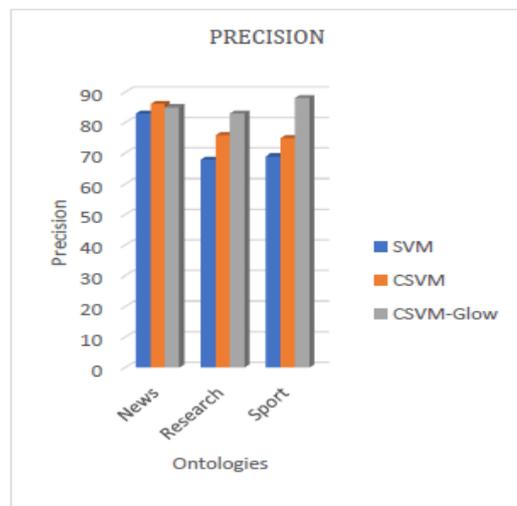


Figure 7: Comparative performance graph showed that the results between SVM, CSVM and CSVM-Glow techniques for the precision parameter in percentage for all news, research and sport ontology.

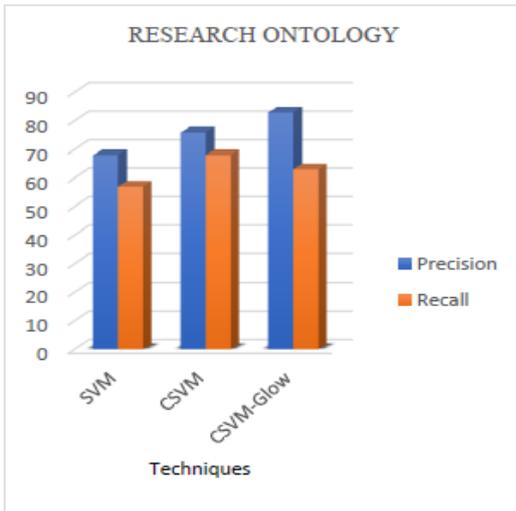


Figure 5: Comparative performance graph showed that the results between SVM, CSVM and CSVM-Glow techniques for precision and recall parameters in percentage for Research ontology.

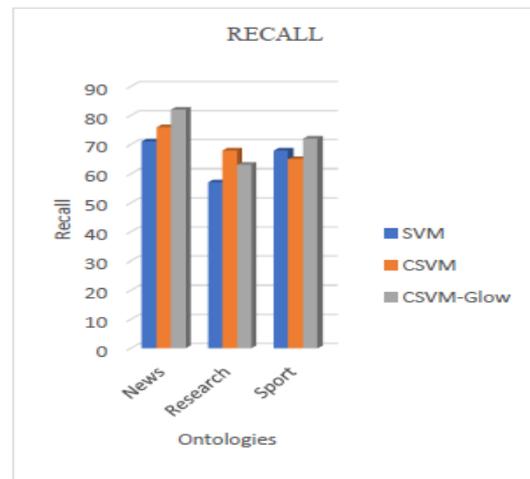


Figure 8: Comparative performance graph showed that the results between SVM, CSVM and CSVM-Glow techniques for recall parameter in percentage for all news, research and sport ontology.

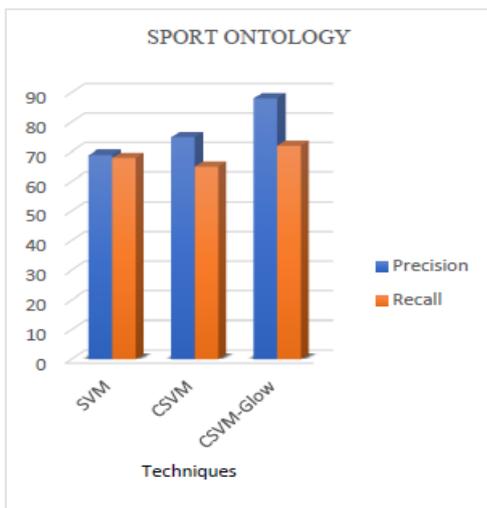


Figure 6: Comparative performance graph showed that the results between SVM, CSVM and CSVM-Glow techniques for precision and recall parameters in percentage for Sport ontology.

V. CONCLUSION

The information extraction process used ensemble-based information extraction. The ensemble-based information extraction used constraints mapping for the domain ontology. The pre-processing of web data used a glowworm optimization algorithm. The ensemble-based information extraction process based on constraints. The ensemble-based classifier used condition for the selection of information extraction in the different lemma. The design information extractor reduces the semantic relation between information extraction and domain ontology. The process of information simplified the web language processing for real-world application. The proposed algorithms tested on various real-time ontology and the result of retrieval suggested some betterment in the sense of hybrid ontology for information extraction and query retrieval.

REFERENCES

1. Farman Ali, Daehan Kwak, Pervez Khan, Shaker Hassan A. E-Sappagh, S. M. Riazul Islam, Daeyoung Park and Kyung-Sup Kwak "Merged Ontology and SVM-Based Information Extraction and Recommendation System for Social Robots", IEEE, 2017, Pp 12364-12379.
2. Farman Ali, Kyung-Sup Kwak and Yong-Gi Kim "Opinion mining based on fuzzy domain ontology and Support Vector Machine: A proposal to automate online review classification", Applied Soft Computing, 2016, Pp 235–250.
3. Zhen-Shu Mi, Ahmad C. Bukhari and Yong-Gi Kim "An Obstacle Recognizing Mechanism for Autonomous Underwater Vehicles Powered by Fuzzy Domain Ontology and Support Vector Machine", Hindawi Publishing Corporation, 2014, Pp 1-11.
4. O. Abuomar, S. Nouranian, R. King, T.M. Ricks and T.E. Lacy "Comprehensive mechanical property classification of vapor-grown carbon nanofiber/vinyl ester nanocomposites using support vector machines", Computational Materials Science, 2015, Pp 316–325.
5. Maciej Zieba, Jakub M. Tomczak, Marek Lubicz and Jerzy Swiatek "Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients", Applied Soft Computing, 2014, Pp 99–108.
6. Zehra Camlica1, H.R. Tizhoosh and Farzad Khalvati "Medical Image Classification via SVM using LBP Features from Saliency-Based Folded Data", IEEE, 2015, Pp 1-5.
7. Ignas Kukenys and Brendan McCane "Classifier Cascades for Support Vector Machines", IEEE, 2008, Pp 1-6.
8. Esraa Elhariri, Nashwa El-Bendary, Mohamed Mostafa M. Fouad, Jan Platos, Aboul Ella Hassanien and Ahmed M.M.Hussein "Multi-class SVM Based Classification Approach for Tomato Ripeness*", Springer, 2014, Pp 175-188.
9. Wenjia Li, Jigang Ge and Guqian Dai "Detecting Malware for Android Platform: An SVM-based Approach", IEEE, 2015, Pp 464-469.
10. Wei Zhou, Junhao Wen, Qingyu Xiong, Min Gao and Jun Zeng "SVM-TIA a shilling attack detection method based on SVM and target item analysis in recommender systems", Neurocomputing, 2016, Pp 197–205.
11. Harsh Arora and Govind Murari Upadhyay "A Framework for the Detection of Suspicious Discussion on Online Forums using Integrated approach of Support Vector Machine and Particle Swarm Optimization", International Journal of Advanced Research in Computer Science, 2017, Pp 1292-1297.
12. Agnieszka Konys "Towards Knowledge Handling in Ontology-Based Information Extraction Systems" Procedia Computer Science, 2018, Pp 2208–2218.
13. Kamran Munir and M. Sheraz Anjum "The use of ontologies for effective knowledge modelling and information retrieval", Applied Computing and Informatics, 2018, Pp 116–126.
14. Diana Maynard, Milena Yankova, Alexandros Kourakis and Antonis Kokossis "Ontology-based information extraction for market monitoring and technology watch*", Content-driven Knowledge Management Tools through Evolving Ontologies, 2015, Pp 1-10.
15. Jie Chen, Chunxia Zhang and Zhendong Niu "A Two-Step Resume Information Extraction Algorithm", Hindawi, 2018, Pp 1-9.
16. Juan F. Sequeda and Daniel P. Miranker "A Pay-As-You-Go Methodology for Ontology-Based Data Access", IEEE, 2017, Pp 92-96.
17. Kamel Nebhi "Ontology-Based Information Extraction from Twitter", Proceedings of the Workshop on Information Extraction and Entity Analytics on Social Media Data, 2012, Pp 17-22.
18. Xiaomeng Su and Jon Atle Gulla "An information retrieval approach to ontology mapping", Data & Knowledge Engineering, 2006, Pp 47-69.
19. Daya C. Wimalasuriya "Ontology-Based Information Extraction", Winter, 2009, Pp 1-37.
20. Apichai Suesatsakulchai, Marut Buranarach and Orasa Tetiwat "Development of Blood Donor Complication Semantic Retrieval System using the Ontology Application Management Framework", International Journal of Applied Engineering Research, 2018, Pp 2361-2367.
21. Rizwan Iqbal, Masrah Azrifah Azmi Murad, Aida Mustapha, Aida Mustapha, Syed Rizwan Ali and Clay Palmeira da Silva "Evaluating Effectiveness of Concept Maps for Ontology Conceptualization: A Quantitative Study", International Conference on Information Retrieval and Knowledge Management, 2018, Pp 119-123.

AUTHORS PROFILE

Ritesh Kumar Shah, has completed his M.Tech(CSE) from RGPV, Bhopal, MP. He is currently pursuing a Ph.D (CSE) from Mewar University, Gagnar, Rajasthan. He has published more than 6 research papers published in International journals and conference. His research area are in Semantic Web, Machine Learning.



Dr. sarvottam Dixit, is an M.E. (Computer Science), Ph.D. (Material Science) from "Agra University" (now called "Dr. B. R. Ambedkar University"), India and has done Post Doctoral Research work in AstroPhysics at TIFR, Mumbai. He is currently working as Professor in Department of Computer Science and Engineering, Mewar university, Chittorgarh (Rajasthan). He is a member of various Computer Societies. He has published more than 35 research papers in reputed International Journals and Conferences. His main research work focuses on Cryptography Algorithms, Machine Learning, Big Data analytics and Computational Intelligence based education.

