Explicit Link Discovery Scheme Optimized with Ontology Mapping using Improved Machine Learning Approach

Sengodan MANI[1]*, Samukutty ANNADURAI[2]

[1]Assistant Professor, Department of Computer Science and Engineering,
Nehru Institute of Engineering and Technology, Coimbatore-641 105, Tamil Nadu, India
maniscovai@gmail.com (*Corresponding Author*)

[2]Advisor & Senior Professor, Department of Computer Science and Engineering,
Hindusthan College of Engineering and Technology, Coimbatore-641 032, Tamil Nadu, India
anna_prof@hindusthan.net

Abstract: The web of data in the field of information and communication technology has been growing steadily in recent years, but there is a severe lack of association between the similar domains. The published web of data often reflects identical types of data that are described in various formats and generated at different locations. Interoperability and heterogeneity problems are generated when accessing such data, which can be solved by combining related ontologies with similarity-matching techniques. The semantic web allows information to be interpreted more meaningfully, providing a description of its contents and services in a machine-readable form called the web of data, which is typically structured in metadata. To process this data, different ontology matching methods are available, such as Silk, Kno Fuss, GLUE, LogMap, AgreementMaker, LIMES, CODI, SERIMI, RiMOM, etc. These approaches focus primarily on the classes of entities and their relationships, not on the principles of each type. In this paper a new model of similarity matching techniques is presented with the purpose of integrating related ontologies. The proposed model includes the entity behaviour and structural information of the ontology classes for the similarity matching process. The paper also includes two machine learning approaches, the first being lexical-based similarity matching using the Threshold-based Support Vector Machine (TSVM), whih is performed with restricted clustering and classification. The second is instance-based similarity matching using the Semantically enhanced Nearest Neighbour method (SeNN), which is employed in order to compare and quantify the semantic enhanced nearest neighbour entities/labels to predict the exact similarities. The final process involves the mapping of two sets of links based on the similarity of domain/ range criteria for accurate results. The proposed approach is compared with the existing state-of-the-art systems and the findings are analysed for precision, and specificity with respect to f-measure values that show better results in comparison with current approaches.

Keywords: Constrained based matching, Threshold-SVM, Semantic enhanced Nearest Neighbour (SeNN), Lexical similarity, Ontology matching, Similarity matching.

1. Introduction

The numerous data sources on the website have played a critical role in many semantic web applications in recent days. In the semantic web, data sources are interpreted as a class of entities with ontology, and their relationships are represented in structured format (Shvaiko & Euzenat, 2008). By combining related data sources one could navigate data sources more easily. So creating a set of binary relations between the vocabularies of two ontologies is the task of aligning ontologies (Yannis & Marco, 2003). Ontologies can be built by using the Ontology Web Language (OWL) (Shvaiko & Euzenat, 2008). It uses the *owl:sameAs* to interlink various definitions, knowledge and instances (Shvaiko & Euzenat, 2011).

The relationship between two entities is based on the concept of linked data principles that has a distinct meaning for various applications (Nentwig et al., 2017). Ontology Alignment is based on semantic similarities between entities in which data can be easily identified by performing machine learning classification. In general the Machine Learning and Deep Learning techniques are used by recent researchers for big data analysis to predict positive and negative groups (Al-Garadi et al., 2019).

Further, the large scale data is also performed using machine learning approach for various applications in semantic web. Pattern analysis is often considered and combined with machine learning techniques for a specific classification of a web of data. In addition to this, the ontology visualization technique in semantic web utilizes source and target data for concept model based relationships (Yuan et al., 2020). For the performance evaluation, the visual dimension of the retrieved data and its characteristics are considered and are categorized normally in a hierarchial manner (Zhang & Lee, 2004). Besides that, the inductive logic programming method for ontology mapping aims to find as many valid

mappings as possible between ontology concepts or entities, particularly for large ones. (Karimi & Kamandi, 2019).

Linked Data Principles has a representation of identity link with a property of RDF type (Shvaiko & Euzenat, 2008). This enables the OWL to create a variety of ontological applications such as DBpedia, Genomes, and FOAF that can be further aligned with several other datasets. In general, the semantic web relativity can enable one to find the concept of entities that map ontological uniqueness (Muhleisen & Jentzsch, 2011).
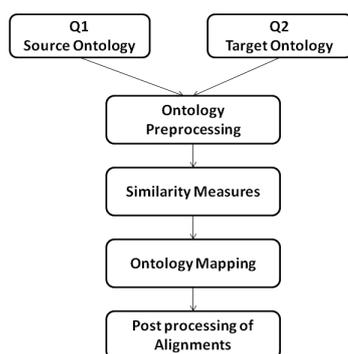
Figure 1. Basic flow of ontology mapping process

Figure 1 illustrates the general flow model of ontology mapping process which gives information about common ontology matching approach. For various real-world applications such as ontology integration, information retrieval and peer-to-peer communication, this approach is applicable (Thiéblin et al., 2020). The generic flow model takes the input as two ontologies, namely Q1 and Q2 that are loaded as in the memory model for preprocessing task. The input ontologies are generally parsed and normalized in preprocessing stage. The similarity of the ontology entities is determined by using the similarity computation. The findings are post-processed in order to avoid any inconsistencies. The validation is also done to map the right alignment in the RDF format that represents the relationship between the source and target entities of underlying ontologies.

The machine learning approach was implemented in the proposed work for better performance. The machine learning approach includes the Threshold Support Vector Machine algorithm (TSVM) and Semanticaly enhanced Nearest Neighbour algorithm (SeNN). TSVM is employed in order

to classify the entity labels and SeNN to classify the final alignments for categorizing the exact mappings. For the analysis and evalution, the information retrieval metrics are considered such as precision, recall and f-measure in order to carry out the performance measures.

This paper is organized as follows. Section 2 includes a literature survey of the related works, while section 3 describes the specific existing methods. Section 4 proposes the novel technique for similarity matching using machine learning approach and section 5 presents the results and discussion part, which compares the performance results with those of the existing approach. Finally, section 6 concludes the proposed methodology and presents future enhancement approaches.

2. Literature Survey

Syntactic and semantic data are the basis of the element-level matching approach and structural-level matching approach in ontology mapping (Nguyen & Conrad, 2015). These are typically based on formal & informal resources, string data, linguistic data, instance data and individual data (Li et al., 2018). The data of meaning, content, structure, element, semantic, syntactic, terminology, and extensional entities is a major goal of the ontology matching approach (Liping et al., 2007).

The visual analytic measure of ontology mapping approach is constructed with VA model and it is proposed as VIS4ML. The CNN and RNN approaches are applied in machine learning based ontology mapping. Various examples are examined with the web of ontology language and the IO entity is applied to the data preprocessing, which reduces the performance by increasing the number of stages in ML (Sacha et al., 2018).

Kaladevi and Geetha (2016) have reviewed various ontology mapping techniques based on semantic web data. This review investigates the applications of semantic-based ontology mapping, particularly in the fields of data integration, data transformation, and information retrieval systems. It compares data between an object and a source's reference point based on concept/technical characteristics. OAEI database is utilized for the variety of applications in ontology mapping. The methodologies used here are typical of semantic

web mining technique, which utilizes structure, instance, terminology, semantic reasoning and other hybrid approaches (Osman et al., 2021).

Heflin & Song (2016) have described that ontology instance mapping model is based on knowledge graph analysis. Interlinked data is mapped and provides the knowledge graphs for identifying ontologies based on the concept of entities. The scaling entities are used to perform large-scale datasets using an instance-based matching approach (Zhang & Haglin, 2016). Semantic web data contains the format of RDF and OWL (Otero-Cerdeira et al., 2015). OAEI benchmark dataset performs data transformation and data deletion using LogMap, Serimi, EPWNG and DisNGram testing tools. The drawback of this method shows up when performing the interlinking of ontologies with multiple selections. A modified version of fuzzy logic was proposed in ontology mapping in real-time XML data (Li et al., 2018). The GLUE is one of the semi-automatic techniques for ontology matching but the major problem of GLUE is that it requires a large number of instances associated with the nodes in taxonomies and these are not available in most ontology mapping cases (Doan et al., 2003). YAM++ is an extensible ontology matching system but the efficiency of the results depends upon the initial stage of matching (Ngo & Bellahsene, 2012).

To detect and repair unsatisfiable classes during the matching process, LogMap implements a sound and highly scalable ontology reasoner as well as a greedy diagnosis algorithm (Jiménez-Ruiz & Grau, 2011). The Wiktionary Matcher is a simple level matcher that uses an online lexical resource in matching process (Paulheim, 2019).

Shao et al. (2016) have described the instance matching approach based on the iterative framework. This approach is performed with RiMOM tool and it uses algorithms with a view to using the knowledge-based approach. The blocking technique is used in the similarity aggregation model to enhance the experimental results of the matching process. The interactive configuration is aligned on the input data and it is pre-processed by blocking strategy. The similarity obtained by the aggregation method is used to

validate the matching model. Based on the queue model, the subject and object are predicted, which allows one to measure the performance score in order to validate the overall methodology and the instance pair is generated and the experimental results of similarity values are validated.

3. Existing Methods

In ontology matching, although the initial stage of simarity comparison is based on syntactic data, it is extended by making use of structural information (Ngomo & Auer, 2011). Further, various machine learning techniques are employed to perform ontology matching based on the applications of web of data (Zhang & Lee, 2004). In addition, WordNet dictionary is widely used for web data integration, which utilizes lexical similarity and structural similarity methods (Varelas et al., 2005). In order to evaluate the ontology matching outcome, OAEI benchmark dataset is used.

There are many current ontology matching systems, such as RiMOM, GLUE, Lily, CIDER, LogMap, AgreementMaker, YAM++, Silk, MapPSO, TaxoMap, etc., which use lexical and structural information for their computation of similarity (Hussain & Srivatsa, 2012). GLUE uses the machine learning classification model for its similarity computation (Doan et al., 2003; Yuan et al., 2019, 2020).

The existing ontology matching system uses the structural based similarity measure, semantic-based similarity measure and lexical based similarity measures for mapping the web of data (Zheng et al., 2010). Most of the existing approaches focus primarily on the classes of entities and their relationships, not on the principles of each type. The literature survey helps one understand the drawbacks of the existing approach and enables one to propose a new ontology matching model by using machine learning approaches of the threshold-based SVM and the semantic enhanced nearest neighbour. The threshold-based SVM is employed for classification process and the semantically enhanced nearest neighbor is employed for the clustering of semantic mining. The proposed work has improved the ontology mapping model, which aligns related ontologies by determining

the explicit relation between two web datasets, using a machine learning approach. This link discovery scheme can be used in a number of web-based applications.

4. Proposed Methods

The proposed ontology mapping approach is performed with the OAEI database. Here, two forms of evaluation tracks such as the anatomy and conference track from the OAEI 2019 background are employed by the Thresold Support Vector machine learning algorithm for mapping ontologies. Similarity vector can be constructed by constrained based and instance-based approach. The value of the assessed ontology elements may vary depending on the weight of an entity. SeNN approach for explicit link discovery improves the performance by adopting the nearest neighbour selection on ontology mapping. In machine learning, validation metrics such as precision and loss can be used to detect overfitting issues (Dietterich, 1995). This can be corrected during the TSVM model's training phase by using the determined class values. By comparing source and target ontologies, the unique web code is generated which is classified by support vector machine learning approach.

results by comparison with LogMap, Agreement Maker and RiMOM IM, etc. (Shao et al., 2016). Semantic web data is structured on the pre-processing stage. The frequent occurrence of similarity value is predicted by domain or range-based model. Here the overall process flow of the proposed ontology mapping is shown in Figure 2. Then the mapping operation is discovered by converting binary classification of similarity vector with TSVM algorithm (Liu et al., 2012). The similarity value is calculated by evaluating the tokenization of each neighbouring node on datasets; otherwise, untokenized frames are applied to the similarity value. If the similarity value is greater than the threshold value, the tokenized similarity value is added to the next neighbor's similarity value. In the proposed research work, two entities are selected to frame the unique ontology using a decentralized approach. Ontology alignment is mapped with constrained based approach and it maps the similarity matrix based on the weight function of ontologies.
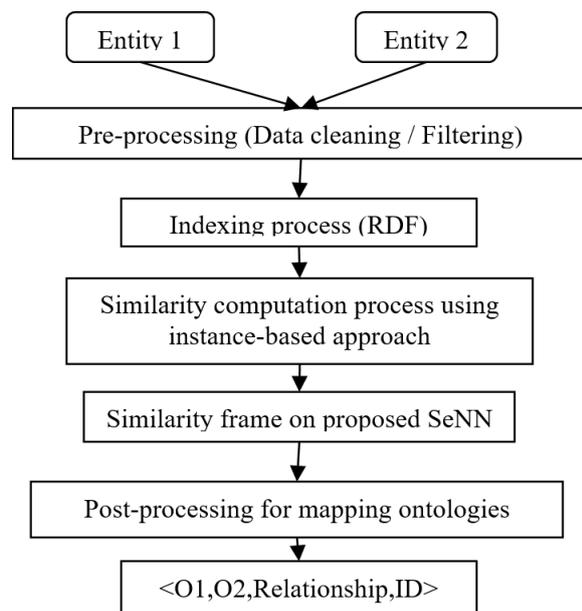
Figure 3. The proposed framework of ontology matching using the proposed SeNN

Figure 2. Process flow of proposed ontology mapping approach using TSVM

To predict the similarity of the entity instances, the semantic enhanced nearest neighbor (SeNN) approach was used here. The proposed work uses the OAEI 2019 dataset, which provides better

The process flow of ontology matching approach is given in Figure 3. It evaluates the unique ontology as a format of <ontology1, ontology2, the relationship between two ontologies (sim (O1, O2)), ID >. Here, the Semantic enhanced Nearest Neighbor (SeNN) is proposed on the instance-based approach and the similarity

matrix is constructed by finding the relationship between two entities using labelled ontologies, which is given in Table 1 below. Following this realisation, semantic web data is used to describe ontologies, which are then mapped using similarity measurements. When mapping from ontology O1 to O2, every entity in ontology O1 finds an appropriate entity in ontology O2 and gives their corresponding relationship. The corresponding relationship is normally determined by the similarity between the two entities (Chua & Goh, 2010).

Table 1. Similarity matrix framing structure

	e2a	e2j	e2m
e1a	Sim(e1a,e2a)	….	Sim(e1a,e2m)
e1i	….	Sim(e1i,e2j)	….
e1n	Sim(e1n,e2a)	….	Sim(e1n,e2m)

O1, O2 are called source ontology and target ontology, respectively. Here $e_{1i} \in O1$, $e_{2i} \in O2$ and $\{e_{1i}\} \underrightarrow{map} \{e_{2i}\} = f$.

If f is null, there are no mapping relationships between e1i and e2j. To find the corresponding relationship between ontologies O1 and O2, entity similarities are computed. The range of values for each element is [0, 1]. The similarity matrix Mk is represented as $M_k = sim(e_{1i}, e_{2j})_{nm}$.

As it is shown in Table 1, where e1i $(1 \leq i \leq n)$ and e2j $(1 \leq j \leq m)$ are elements in ontology O1, and O2, respectively, sim(e1i, e2j) is similarity of e1i and e2j with strategy k (Liu et al., 2012). The initial process of pre-processing includes two tasks, such as data cleaning and filtering for finalizing link configurations. Similarity computation is processed with a matching phase of two ontologies. The dice coefficient is calculated in order to identify the matched instance by varying threshold value.

$$Sim(th) = \frac{2 \,|\, Om \cap Sv \,|}{|\, Om \,| + |\, Sv \,|} \qquad (1)$$

The number of matched outcomes 'Om' is determined by the threshold evaluation and the similarity vector 'Sv' is obtained using TSVM classification. The aim is to improve the similarity performance for adapting more applications related to OAEI dataset. Post-processing technique performed by mapped ontologies and then the performance metric is analyzed for TSVM learning approach.

4.1 OAEI for Ontology Mapping Approach

Ontology Alignment Evaluation Initiative is a coordinating community who evaluates ontology matching tools. The main goal of OAEI is to assess the strength and weakness of the matching process in ontology mapping. OAEI contributes to the improvement of the ontology alignment algorithms and techniques. A large set of semantic and syntactic datasets is loaded to perform the best ontology matching process. Instance matching and web data interlinking are performed based on the OAEI 2019 campaign. Here the multi-strategy mapping approach is employed for mapping similar attributes. The semantic similarity is framed by defining the rules/conditions in order to obtain an unique ontology based on the OAEI database.

4.2 Proposed ML Classification

The proposed research work uses threshold-based support vector machine learning approach for mapped ontology classification based on constrained based approach and the instance-based model utilizes SeNN. Here the proposed Support Vector Machine learning algorithm with an enhanced random selection of threshold-based approach is employed for the classification of extracted data. The Link discovery scheme is performed with a machine learning approach. The trained ontologies are given to the testing process and it provides the performance metrics results such as f-measure, precision, and recall. The mapped data is identified as the expression below.

$$Om = \{(a_1, b_1), (a_2, b_2) \cdots (a_n, b_n)\} \qquad (2)$$

$$Semantic_{Sim}(O_1, O_2) =$$

$$\max \begin{cases} weight(webdata1\,\&\,2), \\ (label1\,\&\,2), \\ (comment1\,\&\,2) \end{cases} \qquad (3)$$

Similarity vector may help one analyze the mapping configurations, which are given by

$$Sv = \{(a_1, b_1), (a_2, b_2) \cdots (a_k, b_k)\} \qquad (4)$$

Both of these above-mentioned equations are represented by the expression below which indicates the total amount of mapping results, which is given by,

$$Ot = |Om \cap Sv| \qquad (5)$$

The threshold-based SVM algorithm uses different threshold values to perform similarity matching. An amount of matched entities are represented with right matching that is 'Ot'; from this, the wrong matching 'w' and non-matching 'n' ontologies are expressed as follows, which helps one evaluate the performance metrics:

Wrong Matching $\{w - Ot\}$ $\qquad (6)$

Non-Matching $\{n - Ot\}$ $\qquad (6.1)$

The feature class is mapped with *RDF: type* of feature code. RDF and CSV files are loaded as the source and target entity and it selects the link type of *owl:sameAs*. Matched rules are selected with a multi-dimensional approach. String similarity measure and numerical analysis are mapped with relationship between each web dataset. The TSVM training algorithm is employed to represent the instance of the given entity and it predicts the categories of anatomy and conference tracks. A clear gap of instance is divided and it maps the high-dimensional matches for obtaining the best match. The training data is mapped with the feature of a task using the feature space vector. SeNN's training data is stored as a sequence pattern, and the testing stage includes an entity pair that tests similarity in order to provide matched ontologies.

In the classification algorithm, the training process extracts the similarity matrix and predicts the actual value and the aggregated value. In the testing process, the training model is extracted in order to obtain the similarity matrix by getting actual value and by comparing actual and system alignment results the performance metrics result is obtained by the classifier. By this proposed model, the similarity measure rectifies the problem of link utilization in memory by the use

of TSVM approach in link discovery scheme. Alignment coherence evaluation is performed in the context of the post-processing approach to generate coherent alignments. For both of these methods, two datasets are processed, and the better similarity results are evaluated.

4.3 Performance Metrics

The proposed research work analyses the performance of improved machine learning approaches that developed a novel technique in SVM and k-Nearest Neighbor by using threshold model and semantic nearest neighbour algorithms, respectively. After matching the similarity values of entities the mapped ontology is creating the unique set of ontology links. The threshold-based support vector machine learning is a supervised machine learning technique, which is applied to the RDF type of web data to perform classification. Similarity computation process uses a semantic relationship between two ontologies and it discovers the link though the similarity matrix formation approach. During the evaluation process the precision, recall, and f-measure score is analysed for improving classification accuracy.

$$Precision = \frac{Ot}{w} \qquad (7)$$

$$Recall = \frac{Ot}{n} \qquad (8)$$

$$Fmeasure = \frac{2TP}{2TP + FN + FP} \qquad (9)$$

Data interlinking process matches the instances by originating schema-based classification. The process involves equivalence mining, coherence, web linkage, and feature configurations. Precision and recall are employed with a view to evaluating the f-measure score for mapped results. The weighted average value of the above analysis helps one find the accuracy of the classifier.

5. Experimental Results and Discussion

Thus the design model of ontology mapping was applied for OAEI test datasets evaluation using threshold-based support vector machine learning approach. The source database is labelled as a, b, c, and so on, and mapping iteration is performed

in order to obtain a similarity matrix using the rule-based constrained match algorithm.

The anatomy and conference track characteristics of the OAEI dataset considered for this study are given in Table 2. The values of precision, recall and f-measure for Wiktionary, LogMap, Proposed TSVM and Proposed SeNN for anatomy and conference track are analysed. The performance of proposed TSVM and Proposed SeNN shows higher values in comparison with Wiktionary and LogMap as it is illustrated in Table 3.

The TSVM model is trained using the class value of the input ontologies being compared, which classifies entity pairs by creating positive and negative classes based on the threshold value. By comparing this, it was found that the overall performance of the proposed ontology mapping technique is better than that of other traditional approaches. Here the improvement achieved through the threshold-based SVM approach and the semantic enhanced nearest neighbor approach contributes to obtaining the results. The graph model of performance values for the anatomy track and conference track is shown in Figures 4 & 5.
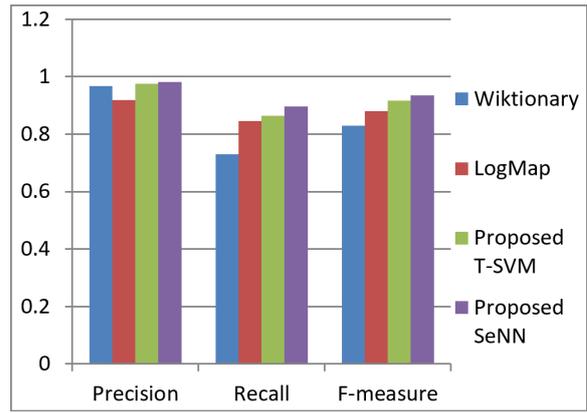
Figure 4. Performance result of Anatomy track
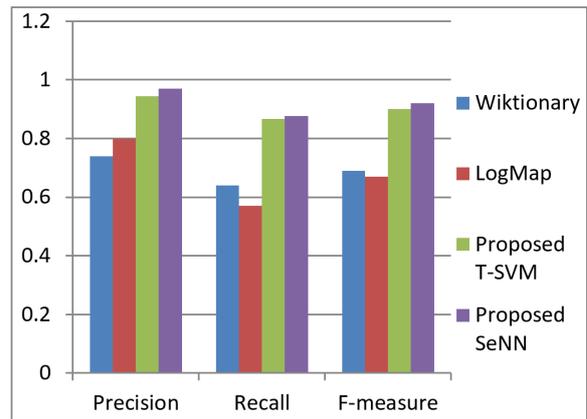
Figure 5. Performance result of Conference track

Table 2. Characteristics of OAEI

Track	Test cases (Track)	Relation	Confidence	Evaluation	Language
Anatomy track	1(11)	=	[0 1]	Open + blind	EN
Conference track	2(9)	=	[0 1]	Open	EN

Table 3. Comparison of results

Methods	Precision	Recall	F-measure
Anatomy track			
Wiktionary	0.968	0.73	0.83
LogMap	0.918	0.846	0.88
Proposed TSVM	0.976	0.864	0.916
Proposed SeNN	0.981	0.896	0.936
Conference track			
Wiktionary	0.74	0.64	0.69
LogMap	0.80	0.57	0.67
Proposed-1TSVM	0.945	0.866	0.90
Proposed -2 SeNN	0.97	0.876	0.92

Figures 4 and 5 illustrate an image comparison of the different methods for ontology mapping, showing that, when compared to Wiktionary and Logmap, both semantic-based analysis and TSVM are capable of mapping two class values effectively, with a precision of over 98 percent.

6. Conclusion and future scope

The ontology matching technique proposed in this paper uses two different approaches to measure the similarities between entities. These are Threshold-based SVM and Semantic enhanced Nearest Neighbor approaches. Both of these supervised learning methods use machine learning's data classification technique to predict related data. Furthermore, both systems learn web data instances and add labels derived from input ontology datasets to the training process. The trained data is used to evaluate the ontology instances in order to predict a similar instance using TSVM and SeNN techniques. Although the

learning and testing phase of the proposed system differ as operation is concerned, both approaches obtain the best results by experimental analysis of the OAEI 2019 dataset. The proposed system is compared with two other common ontology matching systems like Wiktionary and LogMap. The outcomes are evaluated and analyzed in terms of precision, recall and f-measure values. Overall, the proposed system outperforms conventional standard systems in all respects.

A possible future expansion of this work could focus on combining the proposed system with terminological based matching approach and structural-based matching approach to create a collaborative ontology matching system model. Though such a system would largely depend on the estimation of weight calculation due to the variation in different data types under consideration, the recommendation for logistic regression and linear weight assessment could provide promising results in similarity findings.

REFERENCES

Al-Garadi, M. A., Hussain, M. R., Khan, N., Murtaza, G., Nweke, H. F., Ali, I., Mujtaba, G., Chiroma, H., Khattak, H. A. & Gani, A. (2019). Predicting cyberbullying on social media in the big data era using machine learning algorithms: Review of literature and open challenges, *IEEE Access*, *7*, 70701-70718.

Chua, W. W. K. & Goh, A. E. S. (2010). Techniques for discovering correspondences between ontologies, *International Journal of Web and Grid Services*, *6*(3), 213-243.

Dietterich, T. (1995). Overfitting and undercomputing in machine learning, *ACM computing surveys (CSUR)*, *27*(3), 326-327.

Doan, A., Madhavan, J., Dhamankar, R., Domingos, P. & Halevy, A. (2003). Learning to match ontologies on the semantic web, *VLDB journal*, *12*(4), 303-319.

Heflin, J. & Song, D. (2016). Ontology instance linking: towards interlinked knowledge graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, *30*(1), (pp. 4163-4169).

Hussain, M. M. & Srivatsa, S. K. (2012). A study of different ontology matching systems, *International Journal of Computer Applications*, *37*(8), 10-16.

Jiménez-Ruiz, E. & Grau, B. C. (2011). Logmap: Logic-based and scalable ontology matching. In *International Semantic Web Conference* (pp. 273-288). Springer, Berlin, Heidelberg.

Kaladevi, R. & Geetha, G. (2016). Technical review on ontology mapping techniques, *Asian Journal of Information Technology*, *15*(4), 676-688.

Karimi, H. & Kamandi, A. (2019). A learning-based ontology alignment approach using inductive logic programming, *Expert Systems with Applications*, *125*, 412-424.

Li, W., Yan, L., Zhang, F. & Chen, X. (2018). A formal approach of construction fuzzy XML data model based on OWL 2 Ontologies, *IEEE Access*, *6*, 22025-22033.

Liping, Z., Cuangyao, L., Yongguon, L. & Jing, S. (2007). Design of ontology mapping framework and improvement of similarity computation, *Journal of Systems Engineering and Electronics*, *18*(3), 641-645.

Liu, L., Yang, F., Zhang, P., Wu, J. Y. & Hu, L. (2012). SVM-based ontology matching approach, *International Journal of Automation and Computing*, *9*(3), 306-314.

Muhleisen, H., & Jentzsch, A. (2011). Augmenting the web of data using referers. In *Proceedings of Workshop on Linked Data on the Web* (pp. 42-46).

Nentwig, M., Hartung, M., Ngonga Ngomo, A. C. & Rahm, E. (2017). A survey of current link discovery frameworks, *Semantic Web*, *8*(3), 419-436.

Ngo, D. & Bellahsene, Z. (2012). YAM++: A multi-strategy based approach for ontology matching

task. In *International Conference on Knowledge Engineering and Knowledge Management* (pp. 421-425). Springer, Berlin, Heidelberg.

Ngomo, A. C. N. & Auer, S. (2011). Limes - a time-efficient approach for large-scale link discovery on the web of data, *Integration*, *15*(3). DOI: 10.5591/978-1-57735-516-8/IJCAI11-385

Nguyen, T. T. A. & Conrad, S. (2015). Ontology matching using multiple similarity measures. In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K)*, (pp. 603-611). IEEE.

Osman, I., Yahia, S. B. & Diallo, G. (2021). Ontology integration: Approaches and challenging issues, *Information Fusion*, *71*, 38-63.

Otero-Cerdeira, L., Rodríguez-Martínez, F. J. & Gómez-Rodríguez, A. (2015). Ontology matching: A literature review, *Expert Systems with Applications*, *42*(2), 949-971.

Paulheim, H. (2019). Wiktionary Matcher. Ontology Matching. In *Proceedings of the ISWC Workshop* (pp. 181-188).

Sacha, D., Kraus, M., Keim, D. A. & Chen, M. (2018). Vis4ml: An ontology for visual analytics assisted machine learning, *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 385-395.

Shao, C., Hu, L. M., Li, J. Z., Wang, Z. C., Chung, T. & Xia, J. B. (2016). RiMOM-IM: A novel iterative framework for instance matching, *Journal of computer science and technology*, *31*(1), 185-197.

Shvaiko, P. & Euzenat, J. (2008). Ten challenges for ontology matching. In *OTM Confederated International Conferences «On the Move to Meaningful Internet Systems»* (pp. 1164-1182). Springer, Berlin, Heidelberg.

Shvaiko, P. & Euzenat, J. (2011). Ontology matching: state of the art and future challenges, *IEEE Transactions on Knowledge and Data Engineering*, *25*(1), 158-176.

Thiéblin, E., Haemmerlé, O., Hernandez, N. & Trojahn, C. (2020). Survey on complex ontology matching, *Semantic Web*, *11*(4), 689-727.

Varelas, G., Voutsakis, E., Raftopoulou, P., Petrakis, E. G., & Milios, E. E. (2005). Semantic similarity methods in wordnet and their application to information retrieval on the web. In *Proceedings of the 7th annual ACM International Workshop on Web Information and Data Management* (pp. 10-16).

Yannis, K. & Marco, S. (2003). Ontology mapping: the state of the art, *The Knowledge Engineering Review*, *18*(1), 1-31.

Yuan, J., Chen, C., Yang, W., Liu, M., Xia, J. & Liu, S. (2020). A survey of visual analytics techniques for machine learning, *Computational Visual Media*, 1-34.

Yuan, C., Yang, H. & Pan, Y. (2019). Research on Data Link Ontology Mapping Algorithm Based on Bayesian Network Model, *IEEE Access*, *7*, 185698-185709.

Zhang, Q. & Haglin, D. (2016). Semantic similarity between ontologies at different scales, *IEEE-CAA Journal of Automatica Sinica*, *3*(2), 132-140.

Zhang, D. & Lee, W. S. (2004). Web taxonomy integration using support vector machines. In *Proceedings of the 13th International Conference on World Wide Web* (pp. 472-481).

Zheng, C., Shen, Y. P. & Mei, L. I. N. (2010). Ontology mapping based on structures and instances. In *2010 International Conference on Machine Learning and Cybernetics* (pp. 460-465).