# Machine Learning Methods in Ontology Engineering: A Literature Review

Sivaramakrishnan R Guruvayur[1], R.Suchithra[2]
[1,2]*Department of Computer Science, Jain University, Karnataka, India*

**Abstract - Ontology forms a key emerging domain that has a vast potential for improving the organizing, managing and understanding of information. It plays a vital role in the facilitating the access of content, communications, interoperations and in the provision of qualitative and novel services on Semantic Web transformation. The discipline of machine learning (ML) facilitates computers to aid humans in analyzing vast complex repositories of data. The present paper reviews extant literature from the past decade related to the use of machine learning methods in the context of ontology engineering. Certain key ML approaches are identified in this study with general guidelines on the practical uses of ML in varied domains.**
**Keywords- Ontology Engineering, machine learning, supervised learning, unsupervised learning, Support Vector Machines, Decision Trees**

## I. INTRODUCTION

The knowledge-based concept is shifting from a knowledge engineer's laboriously monitored limited standalone device to a distributed self-sufficient knowledge item that is globally accessed by a group of interested persons. Ontologies specifically enable this novel knowledge depiction and, of late, have become a much-studied topic among knowledge engineers. Ontologies offer categorical and formal details of a domain's shared notion that is networked between people disparate and scattered application systems. Developed in Artificial Intelligence, they permit sharing and reusing of knowledge [1].

There are several types of ontologies that are of interest to researchers [2]. These include Domain ontologies which contain knowledge for a specific domain, such as medical, electronics, or an engineering field; Application ontologies which contain information for developing a model of a specific domain by combining method and domain ontologies; Generic ontologies which have validity across many domains; and Representation ontologies which offer entities for representations with no reference to any specific domain, and with no description of what is to be represented (e.g., Frame Ontology).

On the other hand, machine learning (ML) is concerned with developing the methods of extracting inductive patterns from provided data. It delves deep into statistical methods that build the models for data fitting (e.g., regression analysis) and employs highly complex algorithms to gather huge but accurate models that fit the data [3]. These models, in fact, are not even understood by humans any more. Areas which use ML extensively include Forecasting; Adaptivity, e.g., adaptive web-cites, control systems, intelligent agents, profile management; Pattern identification, e.g., text and speech recognition, user behavior modeling, etc.; and Pattern extraction, which generates patterns, which humans can understand.

This present paper endeavors to present a review of extant literature from the past decade pertaining to key applications of ML in the context of ontology engineering. Accordingly, the objectives can be articulated as follows:

1. To identify popularly used machine learning approaches in ontology engineering,
2. To note the trends of studies in the past decade, and
3. To draw conclusions on most popularly used approaches.

The remainder of the paper is organized as follows: section 2 presents the methodology employed for the study, section 3 outlines the ML methods used in ontology engineering in the last decade as evident in existing literature, and section 4 provides the conclusion to the paper.

## II. METHODOLOGY

The following approach was employed for the review of existing literature. First, keywords were identified for the literature search (i.e., Ontology, machine learning, supervised learning, unsupervised learning, Support Vector Machines, Decision Trees). Second, different digital repositories (e.g., Emerald, ScienceDirect, IEEE, Taylor & Francis, Inderscience, Wiley, etc.) were selected for the search based on their relevance and reported quality. The identified papers represented a broad spectrum of research articles and conference papers, mainly peer-reviewed journals. It must be noted that attempts were made using Google search engine to reveal other articles that were not accessible in the online databases. Third, the identified articles were filtered by date of publication to ensure that they were from the last decade. That is, papers published between 2008 and 2018 were considered for the literature review to ensure recentness and appropriateness of research.

An analysis of the ML technique used in the ontologies discussed in the scrutinized papers revealed that the popular techniques were Support Vector Machine (SVM), Decision Trees, Genetic Algorithms, Naïve Bayes (NB), Artificial Neural Networks (ANN), Inductive Logic Program, RF, and Clustering. SVM appeared to be the most popular learning approach (Table 1).

Table 1: Classification of studies according to ML approach

| S No | Authors | Method | Algorithm |
|---|---|---|---|
| 1 | Du et al. [5] | Unsupervised | Clustering |
| 2 | Poon and Domingos [6] | Unsupervised | Clustering |
| 3 | Stocker et al.[9] | Unsupervised | Mixed |
| 4 | Chicco et al.[10] | Supervised | Neural Networks |
| 5 | Qiu et al. [11] | Supervised | Neural Networks |
| 6 | Ngo [12] | Supervised | Decision Trees |
| 7 | Paulheim and Stuckenschmidt [13] | Supervised | Decision Trees |
| 8 | Bijalwan et al. [14] | Supervised | Naïve Bayes (NB) |
| 9 | Luong et al. [15] | Supervised | Support Vector Machine |
| 10 | Ballan et al. [16] | Supervised | Support Vector Machine |
| 11 | Garla and Brandt [17] | Supervised | Support Vector Machine |
| 12 | Xu et al. [18] | Supervised | Support Vector Machine |
| 13 | Záková et al.[19] | Supervised | Genetic Algorithms |
| 14 | Lehmann and Hitzler [20] | Supervised | Inductive Logic Programming |
| 15 | Diamantini et al. [21] | Supervised | Mixed |
| 16 | Rong et al. [22] | Supervised | Random Forests |
| 17 | Belgiu et al. [23] | Supervised | Random Forests |
| 18 | Ongenae et al. [24] | Mixed | Decision Trees |
| 19 | Lehmann and Bühmann [25] | Mixed | Supervised and unsupervised methods |

## III. MACHINE LEARNING METHODS IN ONTOLOGY ENGINEERING

This section outlines the ML methods used in ontology engineering in the last decade as evident from the review of recent literature. It was evident that there are two significant methods of machine learning were utilized by researchers: supervised and unsupervised.

In general, supervised learning is used in an algorithm to understand the mapping function between the input x and the output y, where y = f(x).The purpose is to get a mapping so accurate that output variable y can be predicted with new input data x. As the algorithm learns from a training dataset, it could be seen as a tutor supervising the learning process, thus earning the name supervised learning. The algorithm predicts iteratively from x values in the training dataset and these y values are corrected by the tutor. The learning process terminates on the algorithm's achievement of a satisfactory performance level. Some widely used methods of supervised ML algorithms are Random Forests (RF) for classifying and regression, linear regression, and Support Vector Machines (SVM) for classifying. Most of practical ML applications employ supervised learning [4].

In contrast, in unsupervised learning, input data x is available with mapping output variables absent. The objective here is to create models of the inherent distribution or structure in the data for gathering more information. These methods are referred to as unsupervised learning since there exists no correct solution, and no 'teacher' is available as in supervised learning. These algorithms need to fend for themselves for discovering and presenting the data's inherent structures. Such learning problems could be again classified as association and clustering problems. Often used algorithms for unsupervised learning are apriori algorithms used in problems of association rule learning and k-means in clustering type problems.

The next sub-section reviews studies which deal with applications using unsupervised learning methods.

Applications using Unsupervised Learning Methods

Unsupervised learning methods include clustering analysis and K-means clustering algorithm. Studies using clustering analysis are reviewed in the following sub-section.

### *3.1 Clustering*

Clustering entails the grouping of a collection of items in a manner that items within the same cluster have greater similarity to each other than to items in different clusters. Cluster analysis has found widespread usage in market research for scrutinizing multivariate data originating from surveys. It is also employed by market researchers to create diverse market segments by dividing consumers. This is to aid their understanding of the associations among different customer classes, both current and probable. Other investigations typically facilitated by cluster analysis include developing new products, product positioning studies, pattern recognition, and selection of test markets.
A study by Du et al. [5] proposed a retrieval system for ontologies named OntoSpider for the acquisition of web semantics, with the development of a six-phase method for the ontology extraction from HTML websites by the OntoSpider (Figure 2). This method used information based on the hyperlinks, terms, and tags in the web for semi-automatically extracting data which consisted of preparing, cluster forming (by the use of clustering of instances), recognizing, refining, and revising. In this process, the ontological engineers estimate the parameters and re-examine the concepts, thus playing an important role in assuring the retrieval of a useful ontology. Expected use of this retrieval system encompasses information retrieval from dynamic web pages, updation of knowledge bases, modification of search engines, and searching for key words in blogs. However, this study is limited by its failure to record the impact of semantics, linguistics, and natural language on the quality of the outcomes. Moreover, the effectiveness of the proposed approach could be influenced by the quality of the contents of the pages under consideration. In addition, the study did not explore the manner in which discrete ontologies could be merged into a single universal ontology.

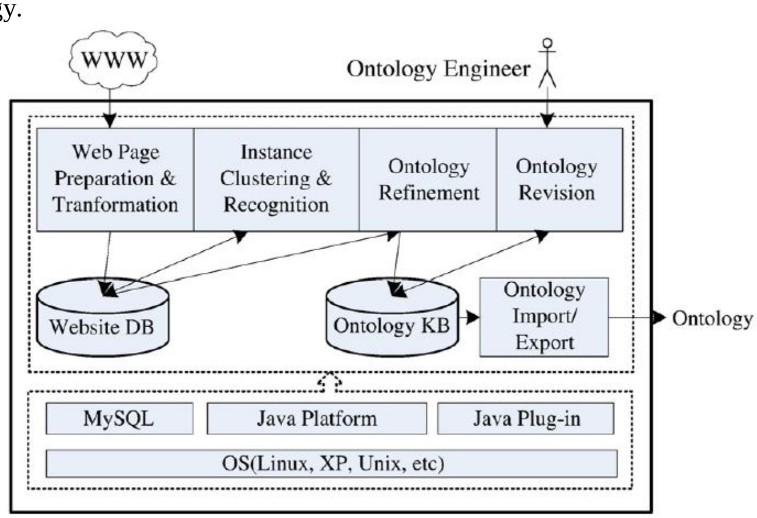

Figure 1: System architecture of OntoSpider [5]

In contrast, Poon and Domingos[6] presented OntoUSP, which is a system for inducing and populating a probabilistic ontology by the use of just input text that is dependency parsed. OntoUSP is built on the unsupervised USP semantic parser using a joint formation of IS-PART and ISA clustering hierarchies of λ-form. Increased general knowledge learning and smoothening for estimation of parameters is enabled by the ISA hierarchy. OntoUSP was evaluated by its use in extracting a knowledge base from abstracts of biomedicine and answering queries. OntoUSP enhances the USP recall by 47% and outperforms earlier methods by a significant margin. However, this study did not explore antonym handling or hierarchical modeling or scrutinizing the learning approach of OntoUSP with the view of extending it to other activities.

*3.2 K-means Clustering Algorithm*

K-means clustering signifies a non-hierarchical approach to arranging items in different clusters [7]. The use case and data can be used as the basis for a user’s definition of the number of clusters. The K-means algorithm “is an algorithm for putting N data points in an I-dimensional space into K clusters. Each cluster is parameterized by a vector m(k) called its mean” [8]. In the K-means algorithm, clustering of data points is accomplished by reducing the total of the sum of squared distances linking the data points and their centroids (i.e., the central point in a data set of data points). The study of Stocker et al. [9] demonstrated the ontological rule learning by the use of numerical measuring and clustering techniques, in particular for anenvironment-based lake ontology with k-means using the average annual data on total nitrogen logged by monitoring centers around lakes. With the learning rules given, rule-based analysis was applied to deduce new information regarding the lakes’ nutrient status. An illustrative example was provided which showed the significance of data on numerical measurement in the learning of environment-

based ontology and the interrelation between ontology engineering and data mining. Nevertheless, this paper highlighted that there was a need to progress beyond theoretical approaches to actual ontology development.
The next sub-section reviews studies which deal with applications using supervised learning methods.
Applications using Supervised Learning Methods
Algorithms using supervised learning use a given group of samples for their predictions. Moreover, the value labels given to the data points are used to search for patterns.

*3.3 Neural Networks*

Artificial Neural Networks(ANNs) are developed utilizing several components of inputs with greater magnitude than computational component possessing typical architectures. A connectionist computation method is used to connect the artificial neurons in categories that utilize mathematical modelling to process information usage. The neurons are kept sensitive for item storage through the ANN. ANNs can be employed for archival of several cases consisting of vectors of high dimensions. Further, the storage can tolerate distortion. An algorithm was developed by Chicco et al. [10] which used deep autoencoder ANNs for aiding curation of annotation databases and predicting new gene functions. This proposed technique need not be restricted to annotation of gene function, but can be utilized in other areas like collaborative filtration of product-recommendation mechanisms. The advantages of the method are: (1) online training ability of autoencoders using big datasets, (2) rapid training option by employing graphics processors, and (3) easy control option of model complexity by the size and number of the concealed layers. The results indicated that autoencoders having two or more concealed layers performed better than single-layered shallow autoencoders, which suggests that deep learning techniques can enhance this application. Again, this paper highlighted that there was a need to progress beyond theoretical approaches to actual ontology development.
Other studies used different approaches to neural networks. For instance, deep neural networks were utilized in a learning method by representation proposed by Qiu et al. [11]. These networks learn the abstract representations of the input entity at a high level. At first, the entity representations are learned in an unsupervised manner, followed by fine-tuning by a supervised method using the training data. The results of experiments indicated that the method could learn significant entity representations from its descriptive data for improved measurement of the entity similarities. The researchers propose to combine external resources (e.g., search engines, Wikipedia) into their method to enhance its performance further. They also propose to perform more experiments for comparison. Thus, it was evident that this approach was still in an experimental phase and had not yet found practical application.

*3.4 Decision Trees*

Decision trees are a graphical representation employing branching methods to depict all probable outcomes of conditional decisions. Decision trees are composed of internal nodes, branches, and leaf nodes. The internal nodes function as attribute tests whereas the tree branches depict the outcomes of the tests. Finally, the leaf nodes function as particular class labels (i.e., decisions resulting after all attributes are calculated). The classification rules are signified by the route to the leaf node from the root of the tree. A study by Ngo [12] proposed a new matching method that combined different methods in ML (Decision Trees), information retrieval and graph matching for enhancing the quality of ontology matching. Nevertheless, efficiency and recall of the method required further enhancement. Moreover, the approach was restrained by inconsistent removing, lack of user interaction, and absence of instance matching. Again, the approach did not appear to have been practically implemented, indicating the grave mismatch between theoretical exploration and practical implementation in this field.
Another study by Paulheim and Stuckenschmidt[13] showed that the approximation of an A-box consistency evaluation reasoner is possible by training ML decision trees classifiers. Decision trees effectively handle the problem with accuracies exceeding 95% and with computational speeds some 50 times faster than any other ontology reasoner. The trees that result are astonishingly small, having 20 or less decision nodes, thus making this method suitable for application in scenarios having minimal computational resources where, both, computation durations and memory are limited. It could be seen that this approach was suitable for scenarios which did not necessitate 100% accuracy which possibly extends its applicability. Nevertheless, the approach appears to have merit from the perspective of reducing the runtime and computational-intensity of ontology reasoning operations.

*3.5 Naïve Bayes*

Naïve Bayes Classifier ML algorithms facilitate classification of documents, web pages, emails or other long text notes. Bijalwan et al. [14] developed a technique (Figure 3) for classifying natural language text wherein from a training documents set having identified categories, it would predict the category of a new document (query).First, the documents were categorized by KNN-based ML method followed by the return of the one that was most relevant. It was concluded that KNN performs with the highest accuracy when compared with Term-Graph and Naïve Bayes. However, a disadvantage for KNN was found to be its higher time complexity.

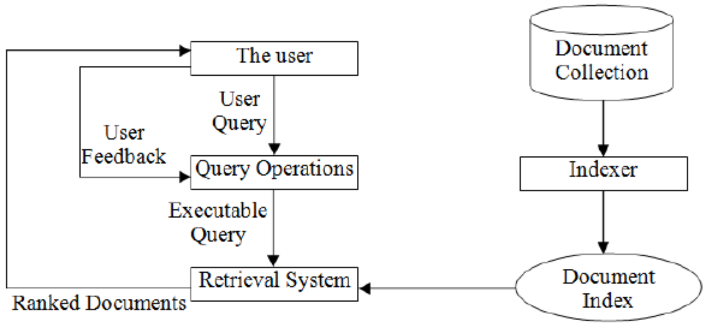


Figure 2: Methodology of document retrieval system [14]

*3.6 Support Vector Machine (SVM)*

As mentioned earlier, SVMs were found to be one of the most popular machine learning techniques utilized in ontology creation. Accordingly, several studies were found to place emphasis on using SVM in ontology generation. For instance, Luong et al. [15] proposed an ontology learning framework that automatically supported tasks of documents retrieval and classification, filtration and extraction of relevant data to enrich the ontology (Figure 4).A focused crawler was developed that enabled document retrieval in the morphology and amphibian domain using A focused crawler was developed that enabled document retrieval in the morphology and amphibian domain using digital library websites. The focus of this work is the assessment of the SVM-based filtration method that involves automated filtration of unrelated documents gathered by the crawler where only those with a high probability of relevance are moved ahead for extracting information. However, this framework is constrained by its lack of applicability to different kinds of documents.

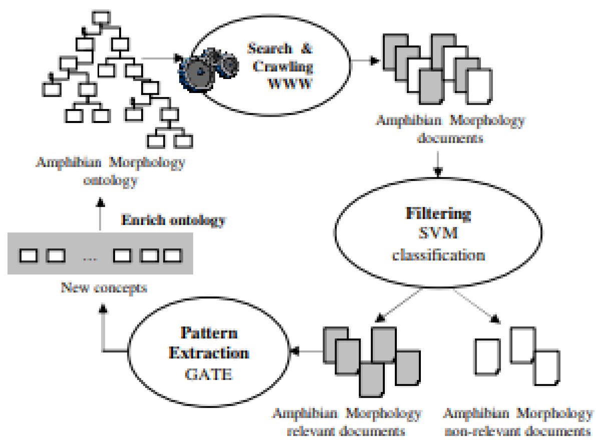


Figure 3: Ontology learning framework architecture [15]

Another method using SVMs was presented by Ballan et al. [16]. This method could be utilized for automatically annotating and retrieving video content from semantic-concept classifiers and ontologies. Semantic linguistic mappings between concepts are automatically determined employing WordNet for defining the schema of the ontology, followed by the linking of concept detectors to the relevant concepts in the ontology. An innovative rule-based technique to annotate semantics of composite events and concepts in videos automatically was suggested. The algorithm learns the Semantic Web Rules Language (SWRL) rules in an automatic fashion, taking in the embedded ontological information. Again this method was limited by its testing only through field trials and professional archivists. Moreover, learning of rules dealing with uncertainty and the usage of fuzzy ontology reasoning required further exploration. A different study by Garla and Brandt [17] proposed an innovative method that ranked features utilizing the domain information encoded in the UMLS taxonomical structure. A semantic similarity metric that is context-dependent was also developed. The performance of the best existing ML-based system was improved by its extension using the proposed methods. These developed methods could enhance the performance of similar ML-based systems that classify clinical texts. Again, the method appeared to be implemented only in an experimental setting.

In their work, Xu et al. [18] described an innovative system that extracted information from narrative summaries of clinical discharge consisting of detailed feature engineering along with ML and rule-based approaches (Figure 5).For the treatment of telegraphic sentences in between ordinary sentences, an approach for a dynamic model switching was proposed. The approach enhanced the concept extractors significantly. Since the system, in medical records, manages telegraphic sentences scattered among ordinary sentences, this concept can be made general and used in other scenarios involving medical records. In classifying assertions, the rule-based classifier output is employed as a statistical classifier feature that succeeds in combining the rule-based method and an ML-based classifier. For smaller classes having limited training data, it performs well. For identifying relations, the paired classifiers architecture significantly improves performance. The novel discriminating features presented in this study were also seen to be effective.

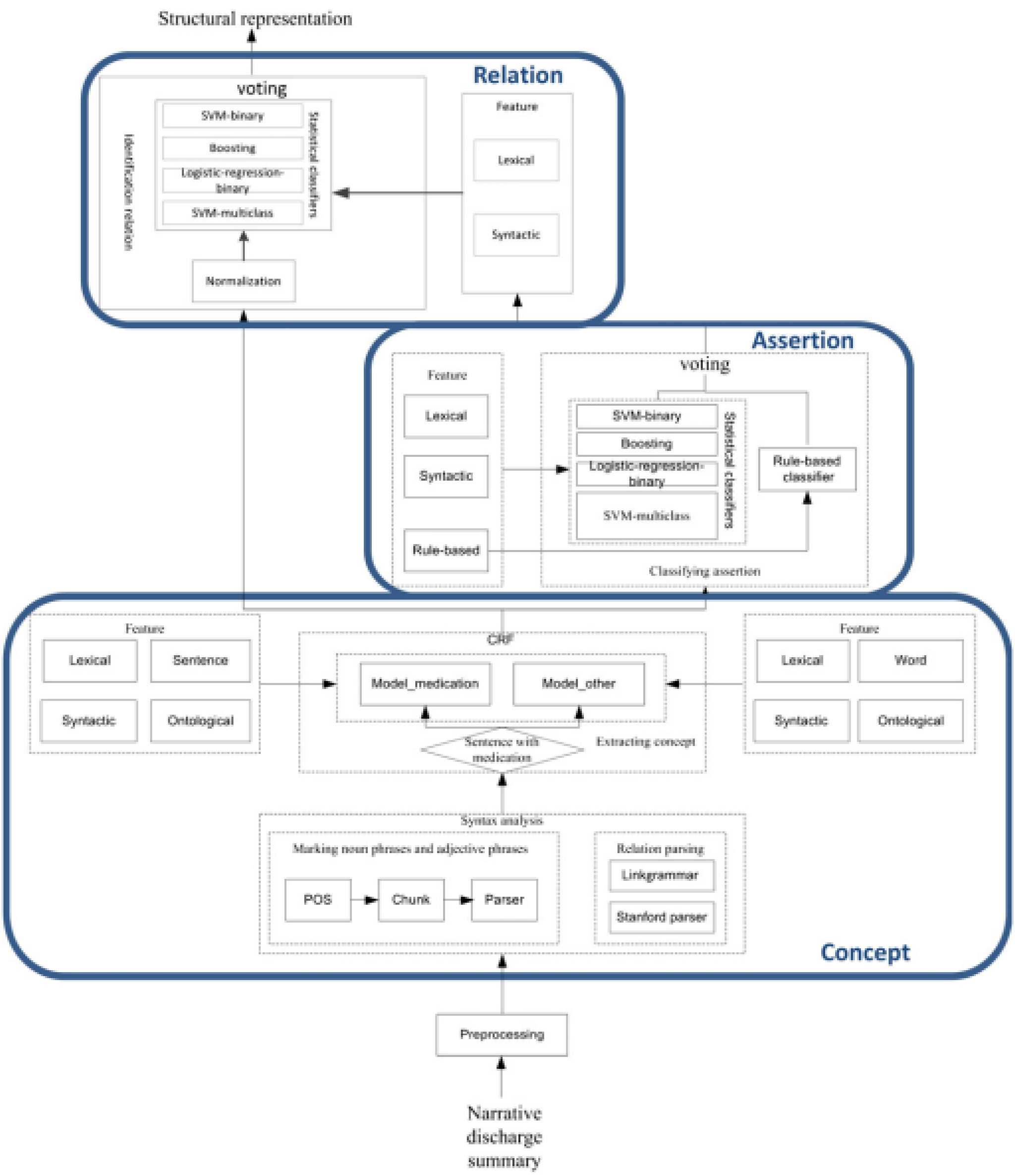


Figure 4: Flow chart of feature engineering/ML system [18]

*3.7 Other approaches*

Other studies were found to use varied machine learning approaches for ontology creation. For instance, a study by Záková et al. [19] used Genetic algorithms in their work on the automated creation of workflows related to knowledge discovery (KD). Their method comprised of two primary parts: (i) defining a formal concept of KD-based data mining and knowledge approaches and (ii) creating a workflow made formal for planning using task detail and domain ontology descriptions. The following forward chaining planning algorithm versions were created: (i) a guideline version demonstrating fitness of the KD ontology for strategizing that used Planning Domain Definition Language (PDDL) algorithm details wherein a method for the conversion of details of data mining to PDDL was generated, and (ii) a second version that placed reasoner-based direct queries for the ontology. The suggested method was verified by two use cases from genomic discovery and advanced engineering. It was seen that the technique offered an encouraging combination of ontological reasoning and planning.

Some other researchers used Inductive Logic Programming as the approach. For instance, Lehmann and Hitzler[20] developed a learning algorithm that worked on refinement operators (ROs) catering to the ALQ description logic (DL) that was inclusive of concrete role support. The algorithm was developed based on a theoretical basis by the identification of probable abstract property combinations that ROs for DLs can have. It was observed that their method was an improvement on other learning methods on DLs, and was comparable with existing ILP methods.

Yet other researchers have used Mixed Algorithms. For example, Diamantini et al. [21] built an ontology (KDDONTO) for data mining by the formalization of the principal components which jointly comprise the inductive bias of an algorithm. A meta-learner can ascertain guidelines for algorithm selection using this ontology by the correlation of the inherent bias of the algorithm with its empirical performance evidence. The developed ontology was designed for supporting meta-learning followed by model selection. They studied the components that built algorithms for revealing common factors beneath their differences. They also identified the inductive-bias components which distinguished every algorithm and its family. The principal components were seen to be (i) the build and features of the produced models, (ii) the employed cost function F that quantified the model's suitability, and (iii) the strategy for optimization used to derive the parameter values of the model which minimizes F.

Another approach used in studies was Random Forests. For example, a method for matching instances that is independent of schema-matching was presented by Rong et al. [22] (Figure 6). The instance matching issue was converted to a classification problem using RF classifiers by the design of an innovative feature vector having similarity metrics of high-level. Appropriate learning models were chosen in accordance with the feature space. The results on IM@OAEI2010 datasets showed that this feature vector matches instances reasonably well, and the method outperformed other existing methods. Further, utilizing the data on existing matches in the Linking Open Data project, new data sources were matched using a transfer learning algorithm.

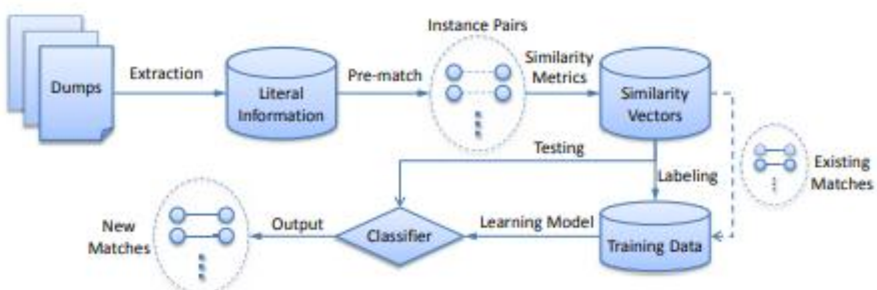


Figure 5: Framework overview of instance matching method [22]

Object-Based Image Analysis (OBIA) techniques were applied by Belgiu et al. [23] extracting building details from Airborne Laser Scanner (ALS) information and studied the ontology-based classification as 'Residential and Small Buildings', 'Industrial Buildings' and 'Apartments' using ML approaches and domain ontologies. These required structures were classified by the exclusive use of the ALS data. RF classifier was used for selection of the relevant characteristics for class predictions. This ontological classification produced fairly good results for the 'Residential and Small Buildings' class (F-measure 97.7%), whereas the other two classes showed lesser accuracy (F-Measure 60% and 51%). This study highlighted the use of ontologies in different domains. The authors propose to overcome certain limitations in the developed ALS data analysis procedure by using laser scanning intensity correction and fine tuning the extraction algorithm. Nevertheless, it is evident that the methodology requires extension before it can be used for the discovery and categorization of different types of buildings in urban settings.

Overall, it could be seen that researchers used varied supervised learning methods in ontology creation. Hence, there did not seem to be many overlaps in the techniques as evident in the studies scrutinized during this literature review. The next sub-section reviews studies which deal with applications using both supervised and unsupervised learning methods.

*3.8 Applications using both Supervised and Unsupervised Methods*

Some studies utilized both supervised and unsupervised machine learning methods. For instance, Ongenae et al. [24] developed a probabilistic, self-learning, ontological framework that allowed adaptation behavior of contextual applications during run-times. The framework (Figure 7) uses the contextual data collected in the ontology for mining patterns and trends in user behavior. The data-mining component uses Decision Trees, Clustering and Bayesian Networks. Such patterns were assigned priority and separated by the association of probabilities. This information and its assigned probabilities then are incorporated within the contextual model and algorithms. The probabilities, finally, are reduced or increased, in accordance with behavioral and contextual data collected about the use of the information learned. The use case presented to demonstrate the framework's applicability showed that accuracy is obtained only if (i) the number of instances in the dataset is at least 1000, and (ii) noise is within 5%.

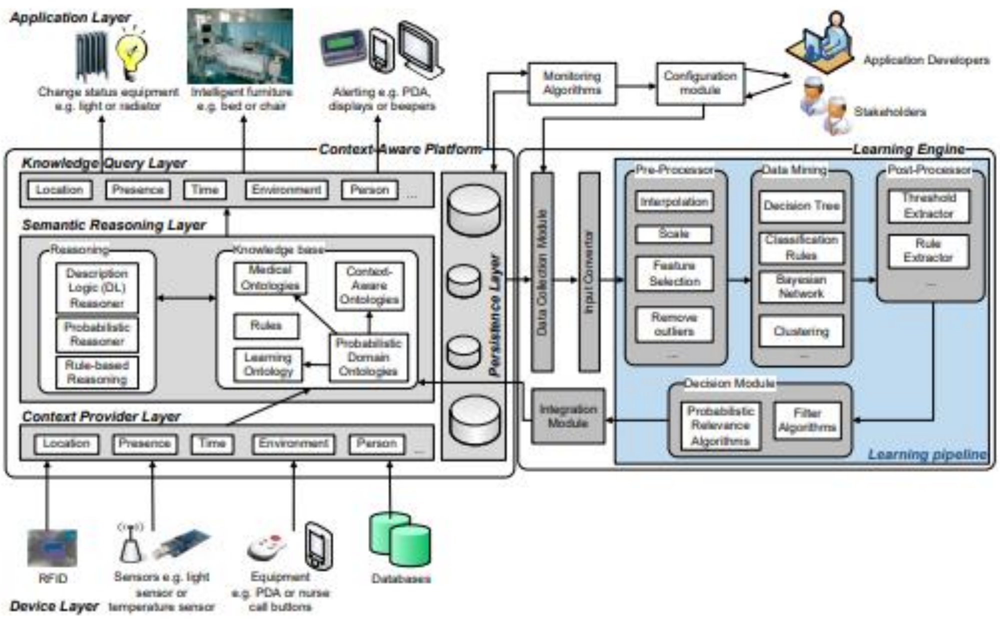


Figure 6: Architecture of the context-aware framework [24]

A combination of ML methods was also used by Lehmann and Bühmann[25] when they developed ORE, a repair and OWL ontology enrichment tool. The cutting-edge methods in debugging of ontologies along with supervised ML form the core of ORE and it has adaptations or extensions that enable smooth working. ORE detects a large number of ontology modeling issues and provides user guidance for the resolution process. Further, the tool permits extension of an ontology by partly-automated supervised ML.

Again, it could be seen that while researchers sometimes used a combination of supervised and unsupervised learning methods in ontology creation, there did not seem to be many overlaps in the techniques as the studies developed ontologies or ontology tools for different purposes. The next section provides the conclusion to this paper.

## IV. CONCLUSION

As stated in the introduction to this paper, the objective of this literature review was to identify popularly used machine learning approaches in ontology engineering, note the trends of studies in the past decade, and draw conclusions on most popularly used approaches.

This study details the use of machine learning algorithms in ontology engineering in the last decade. It could be seen in the past decade that the most popularly used supervised and unsupervised ML techniques used in ontological applications are as follows: SVM, Decision Trees, Genetic Algorithms, NB, ANN, Inductive Logic Program, RF, and Clustering. The most frequently used ML technique was SVM, which finds use in diverse areas such as ontology improvement, healthcare, social analytics, and pharmaceuticals. Thus, it was evident that machine learning finds significant use in ontology generation across different domains and can be inferred to be a vital component in the generation of ontologies.

Table 2 presents the advantages and disadvantages of all the ML approaches presented in this study.

Table 2: Advantages and disadvantages of ML techniques

| Method | Type | Advantages | Disadvantages |
|---|---|---|---|
| Supervised | Support Vector Machine | Performs better with small training examples | Needs apriori information of the observed process distribution<br>Requires data labeling |
| | Decision Trees | Robustness<br>Scalability | Has overfitting issues<br>Has unconstrained nature |
| | Genetic Algorithms | Can handle many parameters<br>Functions well in noisy environments | Uses resources excessively |
| | Naïve Bayes | Simple to use<br>Scalable | Makes unrealistic assumptions of data distribution shape |
| | Neural Networks | Needs no apriori information of the observed process distribution | Has overfitting issues<br>Requires data labeling |
| | Inductive Logic Program | Can learn a wide range of inputs<br>Can include additional background knowledge in the learning problem | Relatively inefficient<br>Limited numerical data handling |
| | Random Forests | Scales well with large datasets<br>Has no overfitting issues | Cannot predict rare outcomes<br>Interpretation is not easy |
| Unsupervised | Clustering | Data groupings can be visually observed | Performs well only when clusters are pre-labeled |

Nevertheless, a paucity of research could be observed with regard to the theoretical foundations of the algorithms while a greater emphasis was placed on the experimental outcomes of approaches devised using these. Furthermore, in most studies, it could be seen that there was no indication of how the extracted data was stored or managed. Moreover, though several of the methods appeared to be automatic there was insufficient detail regarding the automatic generation of an ontology. Further, indications relating to a standard approach to normalize and harmonize the knowledge representation across domains for the purpose of building relevant ontologies were not evident. Additionally, most of the studies described experimental outcomes rather than practical use of their proposed techniques. Moreover, the purposes of the ontology creation and the outcomes were quite diverse which limited the usage of these studies to build the theoretical foundations of new research. It is hoped that these matters will be considered by future researchers.

In addition, the studies scrutinized did not provide much insights with regard to data extraction (i.e., whether from single or from multiple sources);the nature of knowledge acquisition through machine learning algorithms (i.e., automatic or otherwise); automatic selection of attributes and modeling of relationships between the entitiesinvolved; and constant updation or validation of the ontologies created during the course of the study. Hence, it appears that further research is required to investigate these matters and to propose a new approach to automatically generate an ontology which can be directly utilize in diverse real-time AI applications.

It should be noted that the studies reviewed are merely an indicative sample of the research extant in this area. Nonetheless, it is hoped that the cited references cover major ontological application areas, and provide directional pointers to researchers about the different domains using such techniques.

## V. REFERENCES


[1] D. Fensel, Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce, Springer-Verlag, Berlin, 2000.

[2] R. Studer, V. R. Benjamins, D. Fensel, “Knowledge engineering: principles and methods,”Data and Knowledge Engineering,Vol. 25, No. 1, pp. 161-198, 1998.

[3] T. Mitchell, Machine Learning, McGraw Hill, 1997.

[4] J. Brownlee, “Supervised and Unsupervised Machine Learning Algorithms,” 2016 [Online].Available: https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/. [Accessed Dec. 20, 2018].

[5] T. C. Du, F. Li, I. King, “Managing knowledge on the Web–Extracting ontology from HTML Web,”Decision Support Systems, Vol. 47, No. 4, pp. 319-331, 2009.

[6] H. Poon, P. Domingos, “Unsupervised ontology induction from text,”in Proceedings of the 48th annual meeting of the Association for Computational Linguistics,pp. 296-305,Association for Computational Linguistics, 2010.

[7] S. Gopalani, R. Arora, “Comparing apache spark and map reduce with performance analysis using k-means,”International Journal of Computer Applications, Vol. 113, No. 1, 2015.

[8] D. J. MacKay, Information theory, inference and learning algorithms,Cambridge University Press, 2003.

[9] M. Stocker, M.Rönkkö, F. Villa, M. Kolehmainen, "The relevance of measurement data in environmental ontology learning,"in International Symposium on Environmental Software Systems,pp. 445-453, Springer, Berlin, Heidelberg, 2011.
[10] D. Chicco, P. Sadowski, P. Baldi, "Deep autoencoder neural networks for gene ontology annotation predictions,"in Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, pp. 533-540. ACM, 2014.
[11] L. Qiu, J. Yu, Q. Pu, C. Xiang, "Knowledge entity learning and representation for ontology matching based on deep neural networks,"Cluster Computing, Vol. 20, No. 2, pp. 969-977, 2017.
[12] D. H. Ngo,Enhancing ontology matching by using machine learning, graph matching and information retrieval techniques, Ph.D. thesis, Université Montpellier II-Sciences et Techniques du Languedoc, 2012.
[13] H. Paulheim, H.Stuckenschmidt,"Fast approximate a-box consistency checking using machine learning,"in International Semantic Web Conference,pp. 135-150, Springer, Cham., 2016.
[14] V. Bijalwan, V. Kumar, P. Kumari, J. Pascual, "KNN based machine learning approach for text and document mining,"InternationalJournal of Database Theory and Application,Vol.7, No. 1, pp. 61-70, 2014.
[15] H. P. Luong, S. Gauch, Q. Wang, "Ontology-based focused crawling,"in Information, Process, and Knowledge Management, 2009. eKNOW'09. International Conference on,pp. 123-128, IEEE, 2009.
[16] L. Ballan, M. Bertini, A. Del Bimbo, G. Serra, "Video annotation and retrieval using ontologies and rule learning, "IEEE MultiMedia, Vol. 17, No. 4, pp. 80-88, 2010.
[17] V. N. Garla, C. Brandt, "Ontology-guided feature engineering for clinical text classification,"Journal of Biomedical Informatics, Vol. 45, No. 5, pp. 992-998, 2012.
[18] Y. Xu, K. Hong, J.Tsujii, E. I. C. Chang, "Feature engineering combined with machine learning and rule-based methods for structured information extraction from narrative clinical discharge summaries,"Journal of the American Medical Informatics Association, Vol. 19, No. 5, pp. 824-832, 2012.
[19] M. Záková, P. Kremen, F. Zelezny, N. Lavrac, "Automating knowledge discovery workflow composition through ontology-based planning," IEEE Transactions on Automation Science and Engineering, Vol. 8, No. 2, pp. 253-264, 2011.
[20] J. Lehmann, P. Hitzler, Concept learning in description logics using refinement operators,"Machine Learning, Vol. 78, No. 1-2, pp. 203, 2010.
[21] C. Diamantini, D. Potena, E. Storti, "KDDONTO: An ontology for discovery and composition of KDD algorithms," inThird Generation Data Mining: Towards Service-Oriented Knowledge Discovery (SoKD'09),pp. 13-24, 2009.
[22] S. Rong, X. Niu, E. W. Xiang, H. Wang, Q.Yang, Y.Yu, "A machine learning approach for instance matching based on similarity metrics,"in International Semantic Web Conference,pp. 460-475, Springer, Berlin, Heidelberg, 2012.
[23] M. Belgiu, I. Tomljenovic, T. J. Lampoltshammer, T. Blaschke, &B. Höfle,"Ontology-based classification of building types detected from airborne laser scanning data,"Remote Sensing, Vol. 6, No. 2, pp. 1347-1366, 2014.
[24] F. Ongenae, M. Claeys, T. Dupont, W. Kerckhove, P. Verhoeve, T. Dhaene, F. De Turck, "A probabilistic ontology-based platform for self-learning context-aware healthcare applications,"Expert Systems with Applications, Vol. 40, No. 18, pp. 7629-7646. 2013.
[25] J. Lehmann, L.Bühmann, "ORE-a tool for repairing and enriching knowledge bases,"inInternational Semantic Web Conference,pp. 177-193, Springer, Berlin, Heidelberg, 2010.