Inductive Learning for the Semantic Web: What does it buy?

Editor(s): Krzysztof Janowicz, Pennsylvania State University, USA
Solicited review(s): Philipp Cimiano, CITEC, Universität Bielefeld, Germany; Bernardo Cuenca Grau, Oxford University, UK
Open review(s): Pascal Hitzler, Wright State University, USA

Claudia d'Amato * and Nicola Fanizzi ** and Floriana Esposito ***
Department of Computer Science - University of Bari - Italy

Abstract. Nowadays, building ontologies is a time consuming task since they are mainly manually built. This makes hard the full realization of the Semantic Web view. In order to overcome this issue, machine learning techniques, and specifically inductive learning methods, could be fruitfully exploited for learning models from existing Web data. In this paper we survey methods for (semi-)automatically building and enriching ontologies from existing sources of information such as Linked Data, tagged data, social networks, ontologies. In this way, a large amount of ontologies could be quickly available and possibly only refined by the knowledge engineers. Furthermore, inductive *incremental learning* techniques could be adopted to perform reasoning at large scale, for which the deductive approach has showed its limitations. Indeed, incremental methods allow to learn models from samples of data and then to refine/enrich the model when new (samples of) data are available. If on one hand this means to abandon sound and complete reasoning procedures for the advantage of uncertain conclusions, on the other hand this could allow to reason on the entire Web. Besides, the adoption of inductive learning methods could make also possible to dial with the intrinsic uncertainty characterizing the Web, that, for its nature, could have incomplete and/or contradictory information.

Keywords: Ontology Mining, Inductive Learning, Uncertainty

1. Introduction

The Semantic Web (SW) [3] view is grounded on the availability of domain ontologies to be used for semantically annotating resources. Most of the time ontologies are manually built thus resulting in a highly time consuming task that could undermine the full realization of the SW. For this reason several Machine Learning (ML) methods have been exploited to automatize the ontology construction task [33,23,28]. The main focus is on (semi-)automatically building the *terminology* of an ontology while less attention has been dedicated to the enrichment/construction of the *assertional* part, namely the *ontology population problem*, which results in an even more time consuming task.

In last few years, this problem has been tackled by customizing ML methods such as instance based learning [37] and Support Vector Machine (SVM) [40] to Description Logics (DLs) [1] representation that is the theoretical foundation of OWL[1] language which is the standard representation language in the SW. Specifically, the problem is solved by casting the ontology population problem to a classification problem where, for each individual in the ontology, the concepts (classes) to which the individual belongs to have to be determined [8,14,5].

Both methods for building terminology and assertions only marginally dial with another important problem that emerged in the last few years: "how

* E-mail: claudia.damato@di.uniba.it
** E-mail: fanizzi@di.uniba.it
*** E-mail: esposito@di.uniba.it

[1] http://www.w3.org/TR/owl-features/

to manage the inherent uncertainty[2] of the Web"[3]. To face this problem, some proposals have been formulated. They mainly concern with: how to represent uncertain knowledge [30,32,27] and how to reason in presence of uncertain knowledge [31,10,42]. However, they generally assume that a probabilistic and/or fuzzy knowledge base already exists. Inductive learning methods could be used to build probabilistic knowledge bases by learning the probability that: an inclusion axiom, a relationship between two individuals, a concept assertion hold. Indeed, differently from deductive reasoning (generally adopted in the SW context) where, given a set of general axioms, correct and certain conclusions are drawn by the use of a formal proof, inductive reasoning has as input specific examples/data from which a possible/plausible generalization is computed. This generalization is also able to predict the behavior (i.e. the classification) of new and not previously observed examples.

Reasoning on ontological knowledge plays an important role in the SW since this allows to make explicit some implicit information (e.g. concept and role assertions, subsuption relationships). However, in presence of noisy/inconsistent knowledge bases, that could be highly probable in a shared and distributed environment such as the Web, deductive reasoning is no more applicable since it requires correct premises. On the other hand, inductive reasoning is grounded on the generalization of specific examples (assertions in the SW context) rather than correct premises, thus allowing the formulation of conclusions even when inconsistent/noisy knowledge bases are considered.

In this paper, we survey some inductive learning methods specifically focussing on their applicability for solving various *ontology mining* problems. For *ontology mining* we mean all those activities that allow to discover hidden knowledge from ontological knowledge bases (most of the time concept and role assertions are considered), by possibly using only a *sample* of data. The discovered knowledge could be exploited for building/enriching ontologies. Specifically, we envision the applicability of inductive methods for:

- learning new relationships among individuals
- learning probabilistic ontologies
- (semi)-automatizing the ontology population task
- learning probabilistic mapping for the ontology matching task
- refining ontologies
- reasoning on inconsistent/noisy knowledge bases

In the following some these aspects are analyzed. Particularly, in the Sect. 2 an overview of existing ML methods that have been exploited for solving some ontology mining problems is presented. Proposals on how existing inductive learning techniques can be exploited for facilitating the realization of the SW view are presented in Sect. 3. Conclusions are drawn in Sect. 4

2. The present of Ontology Mining

One of the first proposals for automatically building terminologies is the *ontology learning* task [33]. It focuses on learning ontologies (mainly terminologies) from text documents by the use of clustering methods (drawn from Formal Concept Analysis (FCA) [18]) and association rules [22]. Concepts are extracted from documents by the use of Natural Language Processing techniques [34]. Hence, they are clustered to obtain an initial terminology which is further enriched with new relationships (not necessarily taxonomical) by means of association rules. The main limitations of this approach are: 1) the semantic relations among the terms are not fully clear; 2) the expressiveness of the adopted language is less than OWL.

In order to obtain more expressive knowledge bases, different approaches have been set up [26,23,28]. They assume the availability of an initial sketch of ontology that is automatically enriched by adding and/or refining concepts. The problem is solved as an unsupervised learning problem where given a set of positive and negative examples for the concept to learn, namely a set of individuals that are known to be respectively instances of the concept to learn and instances of the negation of the concept to learn, the goal is building a concept description such that all positive examples are instances of it while all negative examples are not instances.

As regards (semi-)automatizing the ontology population task, the problem has been focused by casting it to a classification problem. Given the concepts of an ontology, all individuals are classified with respect to each concept. In [16,8], the *Nearest Neighbor (NN)* approach [37] is adopted. A new instance (individual) is classified by selecting its most similar training ex-

[2]With the term "uncertainty", a variety of aspects are meant such as incompleteness, vagueness, ambiguity.

[3]http://www.w3.org/2005/Incubator/urw3/

amples (existing individuals in the knowledge base) and by assigning it the class (concept) that is majority voted among the training examples. This required to cope with: 1) the *Open World Assumption (OWA)* rather than the usual *Closed World Assumption (CWA)* generally adopted in ML; 2) the non-disjointness of the classes (since an individual can be instance of more that one concept at the same time) while, in the usual ML setting, classes are generally assumed to be disjoint; 3) the availability of new similarity measures to exploit the expressiveness of DLs.

In [12,14,5], a similar approach is adopted. The main difference is given by the use of SVM [40] rather than *NN* to perform the classification. SVM efficiently classifies instances by implicitly mapping, by the use of a kernel function, the training data and the input values in a higher dimensional feature space where instances can be classified by means of a linear classifier. The application of SVM to DLs representation required the definition of suitable kernel functions to cope with the language expressiveness.

A similar underlying idea has been exploited in [2] where FCA [18] has been used for completing both the terminological and the assertional part of an ontology.

Most of these approaches have also been adopted for performing inductive concept retrieval and query answering, namely for determining the set of individuals that are instance of an existing concept or of a concept generated on the fly from the existing concepts and relationships in the ontology. This is done by classifying all individuals in the ontology with respect to the considered concept. The interesting results of using inductive methods have been: 1) a very low error rate; 2) the ability to induce new knowledge, namely new assertions that are not logically derivable. They can be suggested to the knowledge engineer that has only to validate them. Moreover, most of the inductive methods that have been applied to ontological representation (e.g. NN or SVM) have polynomial complexity which would allow to scale on the whole Web.

3. Inductive Learning for the Future of Semantic Web

The adoption of inductive approaches for ontology mining is mainly motivated by the necessity of: a) semi-automatize the mining of the assertional part of an ontology (i.e. the ontology population task);
b) overcoming the limitations showed by deductive reasoning in the SW context [44], namely its inability to: 1) scale on large ontologies; 2) reason on uncertain knowledge; 3) exploit data regularities. On the contrary, induction can be defined as the process of learning from data. In the following, an overview of how some existing inductive learning methods can be exploited for performing several ontology mining tasks is presented.

3.1. Inductive Learning for building ontologies from Folksonomies and Linked Data

A first fruitful usage of inductive approaches is to automatically build ontologies from source of information such as folksonomies and Linked Data [39]. Indeed, besides of the plethora of text documents and Web pages that are used as input for the *ontology leaning* process [19,6], folksonomies and Linked Data are becoming so popular to constitute a non-negligible source of knowledge. We envision the process of learning ontologies from folksonomies and Linked Data as structured in the following the three steps.

1. *Annotated data are clustered* to create meaningful groups. Well known clustering algorithms such as *K-Means*, *DB-SCAN*, *Simulated Annealing* [24] could be used. Clustering methods are generally grounded on the notion of similarity. Given a set of data, the goal of clustering methods is to find clusters that have high intra-cluster similarity and low inter-cluster similarity [37]. Different approaches could be used: hierarchical, partitional or fuzzy. Hierarchical clustering creates a hierarchy of clusters which may be represented in a tree structure called *dendrogram*[4]. The root of the tree consists of a single cluster containing all data, and the leaves correspond to individual data. Partitional clustering determines all clusters at once, generating a flat set of clusters. Both hierarchical and partitional methods usually assume that clusters are disjoint. On the contrary fuzzy clustering methods allow non-disjoint clusters: an instance can belong, with a certain degree of membership, to more than one cluster at the same time. Applying hierarchical (fuzzy) clustering methods (such as *K-Means* algorithm) to Linked Data, a taxonomy is obtained. It could represent a sketch of an ontology that

[4] A dendrogram is a nested grouping of patterns and similarity levels at which grouping changes. The dendrogram could be broken at different levels to yield different clustering of the data.

is populated with the resources to which Linked Data refer to. However, similarity measures that are able to cope with Linked Data representation need to be exploited. Moreover, at the current stage, no intentional concept definitions are available in the sketch of the ontology. In order to avoid this issue, the second step of the proposed process has to be taken into account.

2. *Concept descriptions for the taxonomy can be learnt* by the use of *conceptual clustering* methods [17] whose goal is to give intensional descriptions of the discovered clusters. Most of the conceptual clustering algorithms such as INC, COBWEB, CLUSTER/2 [21,17,36] often exploit generalization operators applied to propositional representations to set up intentional descriptions of the discovered clusters. The application of conceptual clustering methods to Linked Data will necessarily require the definition of new generalization operators that are able to cope with the considered representation. However, at this stage of our learning process, mainly a taxonomy is available. In order to enrich it with new and potentially more expressive concepts and relationships, the third step of the process has to be considered.
3. Some *data mining techniques such as association rules [22] can be used to further discover frequent patterns* both in a single cluster or in the entire data set. These pattern can be seen as positive examples for a concept (or a relation) to learn via a supervised learning process. However, a supervised learning process usually needs also negative example for the concept to learn. The availability of negative examples could be problematic because of the *OWA*. Indeed, differently from the *CWA* (usually adopted in ML) where negative examples are intended as those examples that are not instance of the concept to learn, in the *OWA* generally adopted in the SW context, negative examples should be instance of the negation of the concept to learn[5]. In this situation, where negative examples could be hardly determined, methods for learning from positive (and unlabeled) examples [48,7] only can be exploited.

[5]The problem does not exist if the examples are labelled by an expert as positive and negative examples of a concept (or relationship) to learn. However, this is not really realistic in an open and wide environment such as the Web.

3.2. Class-Imbalance Learning for Concept Retrieval and Ontology Population

As discussed in [8,2,5], inductive learning can be exploited for (semi)-automatizing the ontology population task by casting this problem to a classification problem and by classifying each individual in the ontology with respect to each concept in the ontology itself. The same approach could be adopted for performing inductive concept retrieval and query answering, namely for assessing all individuals that are instances of an existing concept or of a query concept that is built on the fly by composing (for instance via conjunction and/or disjunction) existing concepts. Induced assertions, namely assertions that cannot be logically derived, could be used for enriching the assertional part of an ontology.

However, as it has been experimentally shown [8, 11], this approach could be less reliable when individuals are not homogeneously spread in the ontology, namely when they are mainly instances of a subset of the concepts in the ontology while the remaining concepts have very few instances. In a setting like this, methods such as NN, that performs classification on the ground of the majority voted class among the most similar training examples, would fail. For instance, considering a case in which 97% of training examples belong to a class A and only 3% of them belong to another class, it will be highly probable that most of the time the classification result will be the class A. *Class-imbalance learning methods* [20,29,47] can be exploited to avoid this problem. They are generally used for performing classification in presence of *imbalanced data sets* [20,29,47], namely data sets where the number of examples of one class is much higher than the others. By the use of sampling techniques, class-imbalance learning methods first create a balanced dataset, namely a data set where instances are homogeneously spread among all categories, and then perform the inductive classification task.

3.3. Inductive Learning for Ontology Refinement

Another important task is ontology refinement. Manually performing ontology refinement could turn out to be a very complex task, particularly for large ontologies. In order to (semi-)automatize this task, *learning Decision Trees* methods [37] could be interestingly used for the purpose. Given a set of positive and negative example for a concept to learn, they return a tree from which a concept description is induced.

The application of these methods in the SW context requires: the specification of positive, negative and unlabeled[6] examples (to cope with the OWA) and the exploitation of refinement operators for DL representations [23,28] giving as a output a *Terminological Decision Tree*[7] from which a new concept definition is derived [15]. Hence the ontology can be refined/enriched by adding the new concept or the whole tree, thus introducing a fine granularity level in the concept descriptions (some tentatives in this direction have been presented in [46,45]). Moreover, *Terminological Decision Trees* can be also exploited for classifying individuals with respect to the learnt concept thus having an alternative way for performing inductive concept retrieval and query answering [8,5].

3.4. Inductive Learning for Ontology Evolution

Another interesting problem that can be tackled via inductive reasoning is *ontology evolution*. Indeed ontologies are not static, they evolve over the time, because new concepts are added (TBox evolution) or most of the time because new assertions are added (ABox evolution). Particularly, the ABox evolution could introduce new concepts that are only extensionally defined while their intentional definitions are missing. *Conceptual clustering* algorithms [17] can be crucial for discovering such kind of evolution [13]. Specifically, they can be employed for discovering *concept drift* or *the formation of new emerging concepts* in an ontology. In order to do this, all instances of the ontology are clustered and an overall evaluation of the clusters (called *global decision boundary*) is computed by the use of well known metrics such as Dunn's Index, Silhouette index, generalized medoid [38,4,25]. A new set of instances is considered as a candidate cluster. To determine its nature, namely if it represents a new concept, a concept drift or an already existing concept, the evaluation of the candidate cluster is performed and it is compared with the *global decision boundary*. If this evaluation is lower than the *global decision boundary* than the candidate cluster is assessed as being an existing concept otherwise it is assessed to represent a new/evolving concept. In the latter case, the intentional cluster description (that is a concept description) can be learned and then merged (by the use of the subsuption relationship) in the ontology. Furthermore, methods for tracking cluster transitions could be also exploited [41].

3.5. Incremental Inductive Learning for Scaling on Large Ontologies

The interest in inductive reasoning and inductive learning methods is not only motivated by the fact that they allow to discover concepts and relationships that cannot be deductively derived. The other main reason is given by the limitation that the deductive approach has showed on reasoning at large scale. To cope with this problem *incremental* inductive learning methods [35,43] are particularly suitable. Indeed, these methods do not need the whole set of data. They are able to learn a first model from a sample of the available training examples and then to update the model when new examples are available. This could allow to learn ontologies, for example, by sampling the Web. Specifically, given an initial sample of the Web, a first (set of) ontology (ontologies) is learnt and then continuously updated when new instances are available. Moreover, differently from the deductive approach that cannot be applied to inconsistent knowledge bases, inductive reasoning is able to process data even in presence of inconsistent or noisy knowledge bases [9,8], situation that could be quite common in an open and heterogeneous environment such as the Web.

3.6. Inductive Learning for Building Probabilistic Ontologies

As showed in [11,8], inductive classification can be effectively exploited for performing inductive concept retrieval and query answering. Since the conclusions drawn from inductive reasoning are typically uncertain, this can be explicitly treated, that is the probability of an inductive result (for instance an individual belonging to a certain concept) could be computed. The explicit treatment of the uncertain results gives several advantages: 1) users or applications can have a measure of the reliability of the inductive results; 2) computed probabilities can be exploited for ranking the answers of a query; 3) a new way of formulating queries which include the chance of requiring likely information/event can be considered [44], i.e. a query of kind *finds all persons that live in Italy that are employees and are likely to own a Ferrari* could be treated; 4) probabilistic ontologies can be automatically built.

[6]Because of the *OWA*, for some instances could be not possible to assess if they belong to a certain concept or its negation so the case of unlabeled example has to be considered.

[7]A terminological decision tree is a decision tree from which DL concept description can be learnt.

Particularly, the last point refers to an another interesting open problem in the SW context: how to manage uncertainty. Even if some existing works have tackled the problem [32,27,10,42], mainly they focus on: (a) how to represent uncertain knowledge; (b) how to reason with uncertain knowledge. Almost all of them assume the availability of uncertain/probabilistic knowledge bases. Building probabilistic ontologies could be a task even more hard than building ontologies. The inherent uncertainty of inductive results could be effectively exploited for the purpose. For instance, the classification results for performing inductive concept retrieval can be accompanied by the probability values for which a certain result is true. Such probabilities can be exploited for building probabilistic ontologies by adopting a framework such as the one proposed in [32].

4. Conclusions

The role of inductive reasoning for ontology mining has been analyzed. A summary of the inductive methods currently adopted in ontology mining has been presented, hence a set (of potential new) ontology mining problems have been addressed and proposals for suitable inductive methods, jointly with a brief analysis of the issues to solve, have been done. The applications of inductive methods for learning probabilistic ontologies is considered one of the most challenging and interesting problems. Moreover, methods for learning event probabilities can be also exploited for assessing probabilistic mapping in the ontology matching task.

References

[1] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *Description Logic Handbook*. Cambridge University Press, 2003.

[2] F. Baader, B. Ganter, B. Sertkaya, and U. Sattler. Completing description logic knowledge bases using formal concept analysis. In M. Veloso, editor, *IJCAI 2007, Proc. of the Inte. Joint Conference on Artificial Intelligence*, pages 230–235, 2007.

[3] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American Magazine*, 2001.

[4] J. C. Bezdek and N. R. Nikhil. Cluster validation with generalized dunn's indices. In *Proc. of the Int. Conference on Artificial Neural Networks and Expert Systems, ANNES '95*, page 190. IEEE Computer Society, 1995.

[5] S. Bloehdorn and Y. Sure. Kernel methods for mining instance data in ontologies. In K. Aberer et al., editors, *Proc. of the International Semantic Web Conference*, volume 4825 of *LNCS*, pages 58–71. Springer, 2007.

[6] P. Buitelaar and P. Cimiano, editors. *Ontology Learning and Population: Bridging the Gap between Text and Knowledge*, volume 167. Frontiers in Artificial Intelligence and Applications, 2008.

[7] F. De Comité, F. Denis, R. Gilleron, and F. Letouzey. Positive and unlabeled examples help learning. In *Algorithmic Learning Theory: 10th Int. Conf., ALT'99*, volume 1720 of *LNCS*. Springer, 1999.

[8] C. d'Amato, N. Fanizzi, and F. Esposito. Query answering and ontology population: An inductive approach. In S. Bechhofer et al., editors, *Proc. of the 5th European Semantic Web Conference*, volume 5021 of *LNCS*, pages 288–302. Springer, 2008.

[9] C. d'Amato, N. Fanizzi, B. Fazzinga, G. Gottlob, and T. Lukasiewicz. Combining semantic web search with the power of inductive reasoning. In F. Bobillo et al., editors, *Proc. of the International Workshop on Uncertainty Reasoning for the Semantic Web at ISWC 2009*, volume 527 of *CEUR Workshop Proceedings*, pages 15–26. CEUR-WS.org, 2009.

[10] C. d'Amato, N. Fanizzi, and T. Lukasiewicz. Tractable reasoning with bayesian description logics. In S. Greco and T. Lukasiewicz, editors, *Scalable Uncertainty Management, Second International Conference, SUM 2008*, volume 5291 of *LNCS*, pages 146–159. Springer, 2008.

[11] C. d'Amato, F. Nicola, and F. Esposito. Analogical reasoning in description logics. In *Uncertainty Reasoning for the Semantic Web I: Revised Selected and Invited Papers*, pages 330–347. Springer-Verlag, 2008.

[12] N. Fanizzi and C. d'Amato. A declarative kernel for $\mathcal{ALC}$ concept descriptions. In F. Esposito and et al., editors, *Foundations of Intelligent Systems, 16th International Symposium, ISMIS 2006, Bari, Italy, September 27-29, 2006, Proceedings*, volume 4203 of *LNCS*, pages 322–331. Springer, 2006.

[13] N. Fanizzi, C. d'Amato, and F. Esposito. Conceptual clustering and its application to concept drift and novelty detection. In *Proc. of the Europ. Semantic Web Conference*, volume 5021 of *LNCS*, pages 318–332. Springer, 2008.

[14] N. Fanizzi, C. d'Amato, and F. Esposito. Statistical learning for inductive query answering on owl ontologies. In A. P. Sheth et al., editors, *International Semantic Web Conference*, volume 5318 of *LNCS*, pages 195–212. Springer, 2008.

[15] N. Fanizzi, C. d'Amato, and F. Esposito. Towards the induction of terminological decision trees. In *Proc. of the Symposium On Applied Computing*, volume 2, pages 1424–1428. ACM, 2010.

[16] Nicola Fanizzi, C. d'Amato, and F. Esposito. Instance-based query answering with semantic knowledge bases. In R. Basili and M. T. Pazienza, editors, *AI*IA 2007: Artificial Intelligence and Human-Oriented Computing, 10th Congress of the Italian Association for Artificial Intelligence, Rome, Italy, September 10-13, 2007, Proceedings*, volume 4733 of *LNCS*, pages 254–265. Springer, 2007.

[17] D. H. Fisher. Knowledge acquisition via incremental conceptual clustering. *Machine Learning*, 2:139–172, 1987.

[18] B. Ganter and R. Wille, editors. *Formal Concept Analysis: Mathematical Foundations*. Springer-Verlag, 1998.

[19] M. Grobelnik and D. Mladenić. Knowledge discovery for ontology construction. In *Semantic web technologies: trends and research in ontology-based systems*, pages 9–27. John Wiley and Sons, 2006.

[20] H. Guo and V. L. Herna. Learning from imbalanced data sets with boosting and data generation: the databoost-im approach. *SIGKDD Explor. Newsl.*, 6(1):30–39, 2004.

[21] M. Hadzikadic and D. Y. Y. Yun. Concept formation by incremental conceptual clustering. In *Proceeding of the International Joint Conference on Artificial Intelligence (IJCAI'89)*, pages 831–836, 1989.

[22] J. Hipp, U. Güntzer, and G. Nakhaeizadeh. Algorithms for association rule mining - a general survey and comparison. *SIGKDD Explorations*, 2(2):1–58, 2000.

[23] L. Iannone, I. Palmisano, and N. Fanizzi. An algorithm based on counterfactuals for concept learning in the semantic web. *Appl. Intell.*, 26(2):139–159, 2007.

[24] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Computing Surveys*, 31(3):264–323, 1999.

[25] L. Kaufman and P. J. Rousseeuw, editors. *Finding Groups in Data: an Introduction to Cluster Analysis*. John Wiley and Sons, 1990.

[26] J. Kietz and K. Morik. A polynomial approach to the constructive induction of structural knowledge. *Machine Learning*, 14(1):193–217, 1994.

[27] M. C. A. Klein, P. Mika, and S. Schlobach. Rough description logics for modeling uncertainty in instance unification. In F. Bobillo and et. al., editors, *Proc. of the 3rd ISWC Workshop on Uncertainty Reasoning for the Semantic Web*, volume 327 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2008.

[28] J. Lehmann and P. Hitzler. Concept learning in description logics using refinement operators. *Machine Learning*, 78:203–250, 2010.

[29] X. Liu, J. Wu, and Z. Zhou. Exploratory under-sampling for class-imbalance learning. In *Proc. of the Int. Conf. on Data Mining*, pages 965–969. IEEE Computer Society, 2006.

[30] T. Lukasiewicz. Expressive probabilistic description logics. *Artif. Intell.*, 172(6-7):852–883, 2008.

[31] T. Lukasiewicz. Uncertainty reasoning for the semantic web. In *Web Reasoning and Rule Systems, Int. Conf.*, volume 5837 of *LNCS*, pages 26–39. Springer, 2009.

[32] T. Lukasiewicz and U. Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *J. Web Sem.*, 6(4):291–308, 2008.

[33] A. Maedche and S. Staab. Ontology learning for the semantic web. *IEEE Intelligent Systems*, 16(2):72–79, 2001.

[34] C. D. Manning and H. Schütze, editors. *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.

[35] R. S. Michalski, I. Mozetic, J. Hong, and N. Lavrac. The multi-purpose incremental learning system aq15 and its testing application to three medical domains. In *AAAI*, pages 1041–1047, 1986.

[36] R. S. Michalski and R. E. Stepp. Automated construction of classifications: Conceptual clustering versus numerical taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5:396–410, 1983.

[37] T. Mitchell, editor. *Machine Learning*. McGraw Hill, 1997.

[38] P. Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20(1):53–65, 1987.

[39] N. Shadbolt, T. Berners-Lee, and H. Hall. The semantic web revisited. *IEEE Intelligent Systems*, 21(3):96–101, 2006.

[40] J. Shawe-Taylor and N. Cristianini, editors. *Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.

[41] M. Spiliopoulou, I. Ntoutsi, Y. Theodoridis, and R. Schult. Monic: modeling and monitoring cluster transitions. In *Proc. of the Int Conf. on Knowledge discovery and data mining*, pages 706–711. ACM, 2006.

[42] Umberto Straccia. Reasoning within fuzzy description logics. *J. Artif. Intell. Res. (JAIR)*, 14:137–166, 2001.

[43] N. A. Syed, H. Liu, S. Huan, L. K., and K. Sung. Handling concept drifts in incremental learning with support vector machines. In *Proc. of the Int. Conf. on Knowledge Discovery and Data Mining*, pages 317–321. ACM, 1999.

[44] V. Tresp, M. Bundschus, A. Rettinger, and Y. Huang. Towards machine learning on the semantic web. In P. C G. da Costa et al., editors, *Uncertainty Reasoning for the Semantic Web I: Revised Selected and Invited Papers*, volume 5327 of *LNCS*, pages 282–314. Springer, 2008.

[45] J. Völker, P. Hitzler, and P. Cimiano. Acquisition of owl dl axioms from lexical resources. In *Proceeding of the European Semantic Web Conference (ESWC 2007)*, LNCS, pages 670–685. Springer, 2007.

[46] J. Völker, D. Vrandecic, Y. Sure, and A. Hotho. Learning disjointness. In *Proceeding of the European Semantic Web Conference (ESWC 2007)*, LNCS, pages 175–189. Springer, 2007.

[47] P. Yao. Comparative study on class imbalance learning for credit scoring. *Hybrid Intelligent Systems, International Conference on*, 2:105–107, 2009.

[48] B. Zhang and W. Zuo. Learning from positive and unlabeled examples: A survey. *International Symposium on Information Processing*, 0:650–654, 2008.