

# A Bootstrapping Approach for Semi-Automated Legal Knowledge Extraction and Enrichment

Silvana Castano<sup>1</sup>, Mattia Falduti<sup>1</sup>, Alfio Ferrara<sup>1</sup>, and Stefano Montanelli<sup>1</sup>

Università degli Studi di Milano

Department of Computer Science - Via Celoria, 18 - 20133 Milano

{silvana.castano,mattia.falduti,alfio.ferrara,stefano.montanelli}@unimi.it

**Abstract.** In this paper, we propose a bootstrapping approach for semi-automated legal knowledge extraction. The approach is characterized by the use of a reference legal ontology that is progressively enriched with relevant concepts and related terms extracted from a corpus of legal documents (i.e., Court Decision documents). Supervised, multi-label classification techniques and black-box model explanation techniques are the core components of the bootstrapping approach i) to associate CD documents with appropriate concepts in the ontology and ii) to choose the terms that are decisive for determining the association between a document and a certain ontology concept, respectively. The goal of the proposed approach is to reduce the manual involvement of legal experts as much as possible and to improve the accuracy of document classification, by progressively enriching the term sets associated with ontology concepts. Preliminary experimental results are finally provided to show the contribution of the proposed approach on a corpus of real Court Decision documents.

**Keywords:** legal ontology · legal knowledge extraction · automated Court-Decision analysis

## 1 Introduction

In the legal domain, Court Decisions (CDs) are documents written in natural language where judges give concrete application of rules and concepts that constitute the law, by deciding whether the law has been violated in relation to the facts. Therefore, CDs are a core component of the legal system since a clear and exhaustive understanding of the judge decisions represents a useful support for the activities of all the actors involved in the legal system. However, quantity, complexity, and articulation of CDs are constantly growing. As a result, effectively extracting the judge decisions about a given crime hypothesis from documents related to real trials is becoming increasingly difficult.

---

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). This volume is published and copyrighted by its editors. SEBD 2020, June 21-24, 2020, Villasimius, Italy.

In such a context, techniques and tools for automated extraction of legal knowledge are strongly demanded, to support annotation, analysis, and understanding of legal documents [2, 10]. Semantic Web technologies are usually employed to create legal knowledge bases, namely *legal ontologies*, derived from i) the law, to formally represent the general rules that are relevant/prominent for specific crime hypothesis in the form of legal concepts, and ii) the case-law, to associate legal concepts with relevant law terminology extracted from CDs [12, 13, 16]. However, the discovery of new legal concepts as well as the annotation of legal documents to determine where and how concepts instances are used by judges, are manually performed by legal experts and it is a time-consuming activity, especially when a large corpus of documents is considered [3].

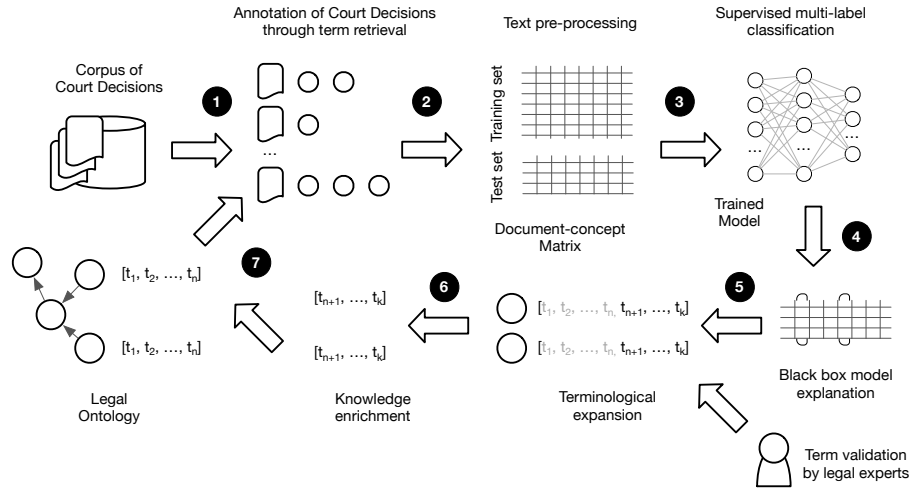
For these reasons, data science approaches are being proposed for automating - as much as possible - the extraction of legal knowledge from textual documents such as Court Decisions. Information retrieval techniques can be employed to detect the occurrence of the terms associated with a concept throughout the documents [7, 17]. In the literature, some contributions are also being proposed in the framework of legal argumentation mining, that is the capability to automatically detect and classify the role of possible argumentative units within a considered legal text [1, 9]. In [15], the authors propose to rely on Natural Language Processing (NLP) and machine learning techniques for mining relevant legal terms from documents. The LUIMA approach characterized by sentence-level annotations and reranking techniques has been also proposed to enforce retrieval over a CD dataset [4]. Moreover, a particularly relevant contribution is provided in [14] about extraction of case law sentences for argumentation of statutory terms, namely terms directly or indirectly defined by the law. However, the accuracy of the above solutions depends on the completeness of the term sets associated with concepts. Due to the variety of terminology adopted by judges in legal documents such as Court Decisions, the construction of accurate and complete term sets to associate with concepts is really hard to obtain.

In this paper, we propose a bootstrapping approach for semi-automated extraction of both terminological and conceptual knowledge in the legal domain. The approach is characterized by the use of a reference legal ontology that is progressively enriched with relevant terms extracted from a corpus of CD documents. Multi-label classification techniques and black-box model explanation techniques are the core components of the bootstrapping approach i) to associate CD documents with appropriate concepts in the ontology and ii) to choose the terms that are decisive for determining the association between a document and a certain ontology concept, respectively. The goal of the proposed approach is twofold. On the one side, the approach aims to reduce the involvement of legal experts as much as possible so that document classification can scale to manage large CD corpora. On the other side, the use of iterative bootstrapping cycles aims to improve the accuracy of document classification, by progressively enriching the term sets associated with ontology concepts.

The paper is organized as follows. In section 2, the proposed bootstrapping approach for semi-automated extraction of terminological and conceptual knowledge in the legal domain is presented. In Section 3, technical details about the adopted machine learning techniques are provided. In Section 4, we present some preliminary results on a real corpus of Court Decision documents. Finally, in Section 5, we give our concluding remarks and we outline our future research issues.

## 2 Semi-automated legal knowledge extraction

Our approach for semi-automated legal knowledge extraction is based on the iterative execution of a *bootstrapping cycle* articulated in a sequence of steps shown in Figure 1. The approach is based on a corpus of Court Decision (CD)



**Fig. 1.** A bootstrapping cycle for semi-automated extraction of terminological and conceptual knowledge

documents and on a reference legal ontology where an initial version of knowledge is provided, both conceptual knowledge and terminological knowledge. We call *conceptual knowledge* the set of *legal concepts* that is formally represented in a reference legal ontology, where concepts are interlinked by semantic relations and associated with a corresponding terminological knowledge. We call *terminological knowledge* the set of natural language terms concretely used in a considered corpus of legal documents (i.e., Court Decisions) to refer to legal concepts. The initial ontology is manually defined by domain experts and it is characterized by a set of legal concepts of interest (conceptual knowledge). A legal concept  $C_i$  in the ontology is associated with an initial *term set*  $T_i^0$  that represents the

relevant terms featuring  $C_i$  that are extracted from the corpus documents since they have been recognized by the experts to be an instance of the concept  $C_i$  (terminological knowledge).

A bootstrapping cycle  $k$  is organized as follows:

**Step (1).** Term retrieval techniques are employed to associate document with relevant ontology concepts. For each CD document  $d$ , the set of associated legal concepts  $\mathcal{C}_d$  is determined as follows:

$$\mathcal{C}_d = \left\{ C_i : \left[ \sum_{t \in T_i} w(t, d) \right] \geq th \right\}$$

where  $w(t, d)$  is the weight of a term  $t$  in the document  $d$  according to standard information retrieval techniques based on tokenization, tf-idf, and PMI (Pointwise Mutual Information) for compound term detection. Moreover,  $th$  is a threshold used to set the minimum cumulative weight of all the terms  $t \in T_i$  that is required for associating a corresponding concept  $C_i$  with the document  $d$ .

**Step (2).** For each document  $d$  in the corpus, a vector-based representation  $\mathbf{d}$  is generated to provide document embedding. In the literature, different techniques can be employed to enforce vector-based document representation, like for example **bag-of-words**, **word2vec**, and **NVSM** (Neural Vector Space Model). In our approach, we choose to rely on **doc2vec** techniques [8]. Basically, **doc2vec** represents an extension of the **word2vec** approach. The **doc2vec** solution has been conceived to overcome the weaknesses of the well-known bag-of-words approach by preserving both ordering and semantics of text-extracted words in the vector representation. In particular, **doc2vec** is based on an unsupervised algorithm that learns fixed-length feature representations from variable-length pieces of texts (e.g., documents). The algorithm represents each document by a dense vector which is trained to predict words in the document. In addition, each document vector  $\mathbf{d}$  is associated with a *concept vector*  $\mathbf{c}_d$ , where each vector dimension denotes a concept  $C_i$  in the legal ontology whose value is set to 1 if  $C_i \in \mathcal{C}_d$ , or it is set to 0 otherwise.

**Step (3).** A multi-label classifier is employed to generate a model that is capable to predict the association of CD documents with legal concepts. In our approach, we employ a 1D Convolutional Neural Network (1D-CNN) with the goal to generalize the terminology of the documents and to enable the association of legal concepts with Court Decisions that actually contain terms other than those already included in the reference legal ontology. For each document  $d$ , the CNN receives the document vector representation  $\mathbf{d}$  as input and it produces the corresponding concept vector representation  $\mathbf{c}_d$  as output. As a result, a classification model  $M$  is generated to map the correspondence between corpus

---

The choice of CNN is due to the positive experimental results we observed in a number of considered case-studies. As a general remark, different kinds of multi-label classifier can be employed for enforcing document classification, like for example random forest and kNN.

documents and legal concepts in the ontology. In particular, by  $C_i \in M(d)$  we denote that the document  $d$  is associated with the legal concept  $C_i$  through the model  $M$ .

**Step (4).** We exploit black-box model explanation techniques in order to select the document features (i.e., terms) that play a major role in determining the decision of the multi-label classifier about the association of concepts with the corpus documents. As a result of Step (4), for a legal concept  $C_i$ , a set  $T_{C_i}$  is generated containing terms that mainly determine the decision of the CNN classifier to associate  $C_i$  with a considered document of the corpus.

**Step (5).** For each concept  $C_i$ , the terms in the set  $T_{C_i} \setminus T_i^k$  are candidate to be exploited for terminological expansion. Legal experts are involved in a validation activity of candidate terms. As a result, for a concept  $C_i$ , the set  $R_i \subseteq (T_{C_i} - T_i^k)$  is defined containing the terms that are relevant for  $C_i$  according to the expert evaluation **Step (6)**. Finally, in **Step (7)**, the terminological knowledge  $T_i^{k+1}$  associated with each concept  $C_i$  is enriched as follows:

$$T_i^{k+1} \equiv T_i^k \cup R_i$$

At the end of Step (7), a new bootstrapping cycle can be enforced. The goal of each bootstrapping cycle is twofold. On the one side, a bootstrapping cycle aims to improve the accuracy of document classification enforced in Step (3). In the first bootstrapping cycle, the accuracy of classification can be low due to the fact that the training set is built by exploiting the terminological knowledge available in the initial version of the legal ontology. As long as the terminological knowledge of the ontology is enriched, the accuracy of the classifier is expected to increase. On the other side, a bootstrapping cycle aims to enrich the terminological knowledge of the legal ontology. The enforcement of new bootstrapping cycles is stopped when the enrichment of the terminological knowledge is terminated, namely when the expert validation (Step (5)) does not generate new terms to insert in the legal ontology.

In the following, more technical details about the black-box model explanation techniques are provided to better emphasize the original contribution of the proposed bootstrapping approach.

### 3 Knowledge enrichment: black-box model explanation and terminology expansion

In a given bootstrapping cycle  $k$ , the goal of black-box model explanation and terminology expansion is to exploit the current version of the legal ontology  $O^k$  and to generate a new version  $O^{k+1}$  where the term sets of THE ontology concepts are enriched with the discovered terminological knowledge. Terminology expansion is based on the multi-label classification model  $M^k$  derived from the annotation of CD documents through  $O^k$ . During the training phase,  $M^k$  learns the function that maps CD documents with terminology of  $O^k$  on the appropriate

legal concepts. In addition, the model also learns to generalize such knowledge, to correctly associate legal concepts with CD documents that actually contain terms other than those included in  $O^k$ . This ability of  $M^k$  depends on two main aspects of the training process. The first one is that CD documents are encoded as vectors using `doc2vec`, thus documents that are semantically similar (but containing different terminology) are encoded as vectors which are “close” in the feature space (i.e., the space of terms). The consequence of this proximity is that the mapping function learned from the model  $M^k$  tends to associate neighboring vectors (i.e., documents) with the same legal concepts. The second aspect is that documents that contain  $O^k$  terms often contain further terms that are also relevant to the legal concepts in the ontology, but which were not discovered/associated in previous bootstrapping cycles. In other terms, the model  $M^k$  implicitly contains the relevant terminology required to map CD documents to legal concepts, even if this terminology is not included in  $O^k$ .

For each concept  $C_i$ , our goal is to detect the set of terms that play a crucial role in determining the classification decision of  $M^k$ , namely the terms that, if deleted from the document, more likely may produce a different classification result. Determining this set of terms is challenging due to the lack of an explicit explanation capable of describing the behavior of  $M^k$ . To this end, we exploit black-box model explanation techniques. Recently, some approaches have been proposed to provide a model explanation at least locally, which means to explain why (i.e., due to which features/terms) a model decides to assign a given class to a certain document [5]. In particular, LIME (Local Interpretable Model-agnostic Explanations) [11] allows to obtain an interpretation of any classifier, by building a local and interpretable model around a prediction. Given a document  $d$ , the idea of LIME is to train an interpretable model using new documents that are uniformly and randomly perturbed copies of  $d$ , located in the proximity of  $d$  by measuring the impact of perturbing each feature on the classification decision. For each term  $t \in d$ , LIME calculates a score  $\eta(t, d)$  that is directly proportional to the relevance of  $t$  in determining the model decision to associate  $d$  with  $C_i$ . Given a concept  $C_i$ , we consider all the documents  $D_{C_i} = \{d : C_i \in M(d)\}$  and all the terminology that is potentially relevant for  $C_i$ , that is:

$$T_{C_i} = \left\{ t : t \in \bigcup_{d \in D_{C_i}} d \right\}$$

Then, we associate each term  $t \in T_{C_i}$  with a degree of relevance  $\eta_{C_i}(t)$  as follows:

$$\eta_{C_i}(t) = \sum_{t \in T_{C_i}} \sum_{d \in D_{C_i}} \eta(t, d)$$

Legal experts are then involved in the validation of terms in  $T_{C_i}$ . A threshold-based mechanism based on the degree of relevance  $\eta_{C_i}(t)$  can be enforced to support the validation activity of experts. In particular, terms with value of  $\eta_C(t)$  higher than the threshold are proposed to the expert for insertion in the legal ontology, while terms with value of  $\eta_C(t)$  lower than the threshold are

proposed to be discarded. As a result of the expert validation, the set  $R_i$  is defined containing the terms that are relevant for the terminological expansion of  $C_i$  so that the new version  $O^{k+1}$  of the legal ontology can be defined.

*Example.* In Figure 2, we show an example of two CD documents,  $d_1$  and  $d_2$  associated with the concept **Drug** in a legal ontology  $O^1$  about the drug criminal legislation (see Figure 3). In our example, the ontology  $O^1$  is implemented by using the Simple Knowledge Organization System (SKOS) [6]. In particular, the legal concepts are implemented as SKOS concepts and they are interconnected through appropriate SKOS relations. For instance, the `skos:related` relation is used to represent a generic positive relationship between two legal concepts, like for example **Drug** and **Criminal Procedure**. For each legal concept (i.e., SKOS concept), a `skos:prefLabel` is defined to denote that a certain term belongs to the term set of the concept. Moreover, a number of `skos:altLabel` are defined to denote the possible alternative terms in the term set of the concept. For instance, a `skos:prefLabel` relation is defined between the **Drug** concept and the **Narcotic Drug** term, while a `skos:altLabel` relation is defined between the **Drug** concept and the **Cannabis** term.

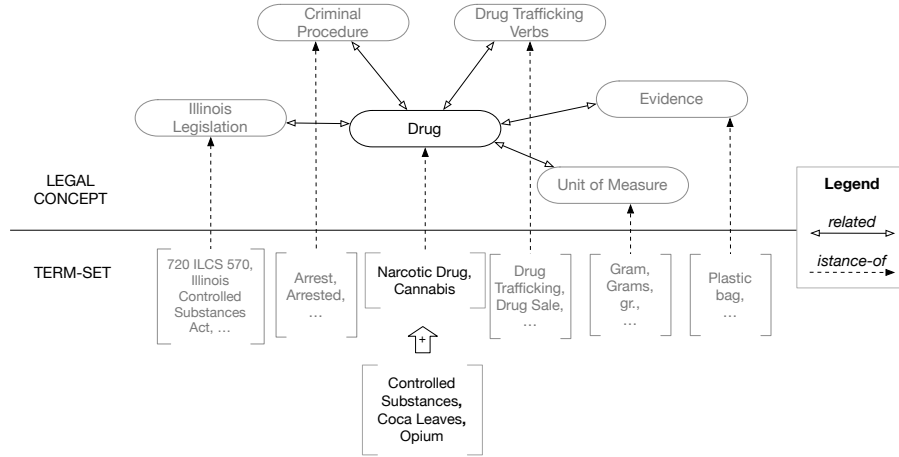
$d_1$ : [...] *Paragraph 14 of section 1 of the same act provides: “**Narcotic Drugs** means coca leaves, opium, **cannabis**, and every substance neither chemically nor physically distinguishable from them.” [...]*

$d_2$ : [...] *Defendant, who was charged by indictment with violation of 402 of the Illinois Controlled Substances Act” [...]*

**Fig. 2.** Example of CD document sentences associated with the legal concept **Drug**

The association of  $d_1$  with the concept is due to the fact that it contains the terms **Narcotic Drug** and **Cannabis** that belong to the term set of the concept **Drug** in the legal ontology. The multi-label classification model  $M^1$  trained on  $d_1$  (and on the other documents contained in the training set) classifies  $d_2$  as a document related to the **Drug** concept. This decision is due to the similarity between the documents  $d_1$  and  $d_2$ , which implies that the two vectors obtained by `doc2vec` are close in the feature/term space.

Through LIME, we detect the terms of  $d_1$  and  $d_2$  that mainly influence the classifier decision. According to LIME, we obtain the following terms for the concept **Drug**: **Narcotic Drug**, **Controlled Substances**, **Cannabis**, **Coca Leaves**, **Opium**. In the list, **Narcotic Drug** and **Cannabis** are already present in the current ontology  $O^1$ , while the others (underlined in Figure 2) are validated by the legal experts. In Figure 3, the validated terms are included in the ontology  $O^1$  to generate a new, enriched ontology  $O^2$  where the term set of the concept **Drug** is properly extended. In the subsequent bootstrapping cycle, the ontology  $O^2$  is exploited to automatically create the training set for the classification model  $M^2$ . Such



**Fig. 3.** Example of the concept Drug in the legal ontology

a training set will include also  $d_2$  since the term *Controlled Substances* has been inserted in the term set of the concept *Drug*. A new round of classification and explanation can be executed to further improve the terminological expansion of ontology concepts and to generate a new ontology version  $O^3$ .

## 4 Preliminary experimental results

The goal of our preliminary evaluation is to assess the i) the capability of discovering new relevant terms about the concepts in the reference legal ontology and ii) the improvement in terms of accuracy of the classification process across two bootstrapping cycles. The experimentation is based on a dataset of around 180,000 Court Decisions of the State of Illinois taken from the Caselaw Access Project (CAP) providing public access to U.S. law (<https://case.law/bulk/download>) digitized from the collection of the Harvard Law Library. For the experiments, we select six concepts from our legal ontology, namely *drug*, *drug trafficking verbs*, *unit of measure*, *illinois legislation*, *criminal procedure*, and *evidence*. Document classification is enforced at the sentence level, which means that legal concepts are associated with each single sentence independently. This way, the 180,000 court decisions correspond to about 14,000,000 documents (i.e., sentences). In the first bootstrapping cycle, the initial version of the legal ontology is characterized by concepts with small term sets (see Table 1). By relying on the term sets in the ontology, we select a subset of 115,993 CD sentences that constitutes the training set of the classification step. According to our annotation techniques, a sentence is associated with a concept  $C_i$  when at least one term belonging to the term set  $T_i$  is contained in the sentence. In Table 2, for each concept considered in the experimentation, we show the number of associated sentences resulting from the annotation step.



**Table 1.** Term sets in the initial legal ontology of the experimentation

Legal Concept	Term Set	Size
drug	narcotic, cocaine, crack, drug, marijuana, heroin, cannabis, lsd, [...]	15
drug trafficking verbs	drug trafficking, drug sale, drug use, narcotics trafficking, [...]	18
unit of measure	gram, grams, gr., pound, kilo, [...]	10
illinois legislation	720 ILCS 570, Illinois Controlled Substances Act, Drug Abuse Control Act, [...]	6
criminal procedure	arrest, arrested, seizure, [...]	7
evidence	plastic bag, plastic bags, paraphernalia	3

**Table 2.** Number of sentences per concept resulting from the annotation step

drug	drug trafficking verbs	unit of measure	illinois legislation	criminal procedure	evidence
25,138	52,103	290	49,994	14,595	2,830

Each document is embedded in a 100-dimension vector using `doc2vec` to obtain a  $115,993 \times 100$  corpus matrix. The model  $M^1$  used to train the classifier is a neural network organized in three layers. Between the input and the output layer, we use a convolution filter activated by ReLU. The  $M^1$  accuracy obtained by cross-validation is 0.77. The model  $M^1$  is then used to perform black-box model explanation and terminology expansion using LIME. For each concept  $C_i$ , we determine a new set of terms  $T_{C_i}$ . A term  $t \in T_{C_i}$  is associated with the degree of relevance  $\eta_{C_i}(t)$ . In the experimentation, a legal expert validated the top-20 terms in the set  $T_{C_i}$  of each concept  $C_i$ . In particular, the expert associated each term  $t$  with a numerical value in  $\{-1, 0, 1\}$ , where  $T^{-1}$  denotes the set of terms that were not in the ontology  $O^1$  and that are not relevant for the concept  $C_i$ ;  $T^0$  denotes the set of terms that were in  $O^1$  (and thus have been already validated as relevant);  $T^1$  denotes the set of terms that were not in  $O^1$  but that are relevant for the concept  $C_i$ .

An overview of the results of terminological expansion is shown in Table 3.

The number of relevant terms retrieved in the terminological expansion (i.e., terms in  $T^0$  or  $T^1$ ) is equal to the 83% of the total number of new terms validated by the expert ( $T_{C_i}$ ). The 34% of those terms was not in the term sets of the initial ontology  $O^1$ . As expected, the increment of new relevant terms is higher for the concepts that were associated with small term sets, such as *illinois legislation*, *criminal procedure*, and *evidence*. The number of irrelevant terms  $T^{-1}$  is limited with the exception of the concept *evidence*, because the criminal evidences usually consist in common objects that are used in a criminal context. These objects are thus associated with a generic terminology (e.g., garbage, suitcase) that cannot be associated per se to an evidence according to the legal expert. The new relevant terms are finally included in the new version  $O^2$  of the ontology that is used to automatically create a new training set for a second bootstrapping cycle. The new training set consists of 158,398 CD sentences (+37% with respect to the

**Table 3.** Results of the black-box model explanation and terminology expansion step

	drug	drug trafficking	unit of	illinois	criminal	evidence	total
	verbs		measure	legislation	procedure		
$ T_i^0 $	15	18	10	6	7	3	<b>59</b>
$ T_{C_i} $	20	20	20	20	20	20	<b>120</b>
$ T^{-1} $	1	0	3	1	6	10	<b>21</b>
$ T^0 $	14	17	10	6	6	3	<b>65</b>
$ T^1 $	5	3	7	13	8	7	<b>34</b>
$\frac{( T^0 + T^1 )}{ T_{C_i} }$	0.95	1.0	0.85	0.95	0.75	0.4	<b>0.83</b>
$\frac{ T^1 }{( T^0 + T^1 )}$	0.26	0.15	0.42	0.68	0.57	0.7	<b>0.34</b>

first execution). In particular, the main increment of sentences is related to the concepts unit of measure (from 290 to 7,241 sentences) and evidence (from 2,830 to 33,417 sentences). These sentences are then used to train a new model  $M^2$  using the same neural network architecture of  $M^1$  and to enforce the execution of the knowledge enrichment steps. Finally, the accuracy of  $M^2$  obtained by cross-validation is 0.81 (+5.2%).

## 5 Concluding remarks

In this paper, we propose a bootstrapping approach for semi-automated legal knowledge extraction. Technical details about the use of multi-label classification techniques and black-box model explanation techniques are provided to show how we associate corpus documents with appropriate concepts in a reference ontology, and how we choose the terms that are decisive for determining the association between a document and a certain ontology concept, respectively. Preliminary results on a corpus of Court Decision documents are discussed to highlight the contribution of our proposed approach in real scenarios. Future work are about the extension of preliminary experiments on a larger corpus of Court Decision documents, and the comparison of obtained results by adopting different techniques for document annotation/embedding, document classification, and black-box model explanation.

## References

1. Ashley, K.D.: Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age. Cambridge University Press (2017)
2. Castano, S., Falduti, M., Ferrara, A., Montanelli, S.: Crime Knowledge Extraction: An Ontology-Driven Approach for Detecting Abstract Terms in Case Law Decisions. In: Proc. of the 17th Int. Conf. on Artificial Intelligence and Law (2019)
3. Grabmair, M., Ashley, K.D.: Facilitating Case Comparison Using Value Judgments and Intermediate Legal Concepts. In: Proc. of the 13th Int. Conference on Artificial Intelligence and Law. pp. 161–170. ACM (2011)

4. Grabmair, M., Ashley, K.D., Chen, R., Sureshkumar, P., Wang, C., Nyberg, E., Walker, V.R.: Introducing LUIMA: an Experiment in Legal Conceptual Retrieval of Vaccine Injury Decisions Using a UIMA Type System and Tools. In: Proc. of the 15th Int. Conference on Artificial Intelligence and Law. pp. 69–78. ACM (2015)
5. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D.: A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys (CSUR)* **51**(5), 1–42 (2018)
6. Isaac, A., Summers, E.: SKOS Simple Knowledge Organization System Primer. Tech. rep., Working Group Note, W3C (2009)
7. Lame, G.: Law and the Semantic Web: Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications, chap. Using NLP Techniques to Identify Legal Ontology Components: Concepts and Relations, pp. 169–184. Springer Berlin Heidelberg (2005)
8. Le, Q., Mikolov, T.: Distributed Representations of Sentences and Documents. In: Proc. of the 31st Int. Conference on Machine Learning. Beijing, China (2014)
9. Nazarenko, A., Wyner, A.: Legal NLP Introduction. *Traitement Automatique des Langues* **58**(2), 7–19 (2017)
10. Palmirani, M.: Legislative XML for the Semantic Web, chap. Legislative Change Management with Akoma-Ntoso, pp. 101–130. Springer (2011)
11. Ribeiro, M.T., Singh, S., Guestrin, C.: “Why Should I Trust You?” Explaining the Predictions of Any Classifier. In: Proc. of the 22nd ACM SIGKDD Int. Conference on Knowledge Discovery and Data Mining. pp. 1135–1144 (2016)
12. Saias, J., Quaresma, P.: Law and the Semantic Web, chap. A Methodology to Create Legal Ontologies in a Logic Programming Information Retrieval System, pp. 185–200. Springer (2005)
13. Sartor, G., Casanovas, P., Biasiotti, M., Fernández-Barrera, M.: Approaches to Legal Ontologies: Theories, Domains, Methodologies, vol. 1. Springer Science & Business Media (2010)
14. Savelka, J., Ashley, K.D.: Extracting Case Law Sentences for Argumentation about the Meaning of Statutory Terms. In: Proc. of the 3rd Int. Workshop on Argument Mining. pp. 50–59 (2016)
15. Savelka, J., Grabmair, M., Ashley, K.D.: Mining Information from Statutory Texts in Multi-Jurisdictional Settings. In: Proc. of the Int. Conference on Legal Knowledge and Information Systems. pp. 133–142. IOS Press (2014)
16. Tiscornia, D.: The LOIS project: Lexical Ontologies for Legal Information Sharing. In: Proc. of the V Legislative XML Workshop. pp. 189–204 (2006)
17. Wagh, R.S.: Knowledge Discovery from Legal Documents Dataset Using Text Mining Techniques. *International Journal of Computer Applications* **66**(23) (2013)