ISSN: 0974-8571 Vol.14 No.1 June, 2022

International Journal of Computational Intelligence in Control

Acquiring Rule-Based Domain Specific Features for Efficient Semantic Similarity Determination in the Judicial Literature

Muhammad Faheem Khan[1,*], Aurangzeb Khan[1],Khairullah Khan[1], Said Khalid Shah[1], Nasir Gul[1], Irfan Ullah Khan[2]

[1] Department of Computer Science, University of Science & Technology, Bannu, 28100 Pakistan

[2] Department of Education and Research, University of LakkiMarwat, 28420 Pakistan

[*] Correspondence: faheemphc@gmail.com

Abstract: Decision making in the legal domain is often a complex task that involves gathering a large amount of information, analyzing issues, legislation and precedents, as well as evaluating feasible options. A lot of work has been done in the field of semantic similarity with multiple approaches in general-purpose domains. The role of semantic similarity is very important in the development of specialized search engine in the Judicial literature. In this paper, we try to acquire the domain specific features to prepare an input vector for efficient semantic similarity determination. We adopt a rule-based approach to acquire domain specific features by using WordNet (general purpose lexicon) and Black Law Dictionary (domain specific judicial dictionary). The methodology is based on the alignment of features in the content words list between the WordNet (synsets) and the definition described in Black Law, considering the lexical forms in both resources to prepare a domain specific input vector for efficient semantic similarity determination. The results obtained from this rule-based approach by combining the two resources makes it possible to acquire the domain specific features and prepare the input vector for further semantic similarity determination which confirmed our hypothesis. The proposed method is evaluated with 6 different categories of data for the judicial literature i.e. Criminal, Civil, Revenue, Services, Constitutional, Corporate. We improved the domain specific features extraction process with an average of 9.28%. As compared to other methods, our proposed rule-based technique achieved an accuracy of 90.2%. (Precision, 0.92, recall, 0.91 and F-measure, 0.90).

Keywords: Semantic similarity; wordnet; black law; judicial literature; input vector;

1. Introduction

Semantic similarity measures has significant importance in many fields of computer applications such as Natural Language Processing (NLP), text summarization, information retrieval and educational system. Mathematically, semantic similarity can be determined by finding the distance between terms or concepts using ontologies. It can also be estimated statistically to correlate the textual context and words from appropriate text corpus such as vector space model [1]. In information retrieval, similarity measure ranks score between text and query in the corpus. The task in natural languages is very challenging when words are same in meaning but not lexicographically [2] Semantic similarity is being used in the major area of NLP such as machine translation, Natural language understanding and Sentiment analysis. Deep learning method is the most popular and accurate way to measure the semantic similarity between two texts [3]. Geoinformatics is used to find the similarity between the concepts in geographics ontologies [4] and tags in OpenStreetMap. [5].

In computational linguistics, several metrics use WordNet and SentiWordNet for calculating semantic similarity [6]. Feature-based method is one of the popular methods which determine similarity between two words/concepts depending on the selected features which is also called "noun" and "is-a" relations-based methods. [7]. Sentiment analysis plays an important role to represent the semantic distance which gives numerical similarity scores to words [8]. If two objects are semantically related it represents semantic relatedness, which is inverse of semantic distance [9]. Semantic similarity can be measured using various methods: (1) Lexicon based (2) Corpus based (3) Knowledge based. Lexicon based approach is mainly used in sentiment analysis with corpus and dictionary approaches using lexical resources such as WordNet or SentiWordNet[10].

Corpus-based methods follow distributional theory for semantic similarity, which is based on word associations learnt from large text. In this method, words are considered more similar if their context is similar or used frequently together. The computation of this method is based on statistics of word co-occurrences or word distributions [11].

There are various count-based methods as per different computational models such as Word2Vec [12], Normalized Google Distance [13] or Pointwise Mutual information [14]. Count-based methods construct word-word matrix, where probabilistic model is directly applied to the word co-occurrences such as matrix factorization [15] and dimension reduction [16]. Word2Vecis a word-embedding tool based on predictive-based method, which directly learns dense vectors by predicting a word from the context of words at its surrounding. Word2Vec is computationally efficient for larger corpus [12].

In knowledge-based method semantic similarity is based on ontology. In this method two words can be assumed more similar if they are positioned closer to each other in a given ontology. The lexical resource WordNet [6] is used as background ontology. The knowledge-based method improves the semantic similarity between the words by encoding the information of human defined hierarchical relations and synonymous words in sysnset[17].

The recent advancement of deep learning explored many research areas in machine learning field such as medical imaging [18, 19], object classification [20, 21], visual surveillance [22, 23], signature / text [24–26].

To deal with these large amounts of text, knowledge-based systems are considered best-performing systems. Knowledge is unlimited and it is not possible for knowledge bases to store all the knowledge for all domain. For many applications based on knowledge bases, an extensive domain specific knowledge is necessary in addition to machine-readable dictionaries built for general-purpose knowledge.

1.2 Need of Acquiring Domain Specific Features

There are different machine learning methods used to predict semantic similarity between textual data. This work investigated the input vector of LSTM model. The LSTM model reads textual data in word-vectors and employs hidden state as a vector representation [3]. To determine the semantic similarity between input text, the domain knowledge is very important. Many English words have different context in the Judicial Domain from the general-purpose domain. To determine accurate semantic similarity between text in the judicial domain, it is necessary to search the meaning and context of that particular word in the judicial dictionary.

The goal of this work is to develop an efficient model that prepares and update input vector of LSTM model for further utilization in semantic similarity determination process.

The study aimed to extract domain specific features from the Judicial dictionary and update the feature list using a Rule-based approach. The LSTM was unable to determine the semantic similarity properly as per the context of judicial domain, and so, judicial dictionary (Black Law) was utilized to accurately capture the context and meaning of feature in the feature list and thus update the feature list with new acquired feature. The process involves domain specific features enriched with general purpose lexicon (WordNet) and Judicial dictionary (Black Law).

1.3 Our Contribution

Following are the contributions of this paper:

1. Rule-based model is developed to prepare the domain specific input vector for LSTM model.
2. Utilize the Black Law dictionary for extraction of domain specific features in the Judicial Domain.
3. Update the features List based on defined rules utilizing Black Law dictionary and WordNet Synset.
4. To utilize the proposed model for acquiring domain specific features and prepare input vector and validating the effectiveness of proposed method by obtaining the domain specific features.

The rest of the paper is organized as follows. In Section 2, the related work comprising of semantic similarity and acquiring domain specific knowledge is discussed. In Section 3, we discussed the proposed model. In Section 4 the experimental setup is explained. Finally, section 5 gives conclusion.

2. Related Work

In this section, different machine learning methods have been reviewed that are used for acquiring domain specific knowledge from text. The works in [27] presented a method that acquire domain specific knowledge from text by extending WordNet. The concepts learned in this method were classified into an ontology and selected for financial domain using three seed concepts: interest rate, inflation and stock market. The proposed method pursuing the knowledge that obtain the concepts including some new concept that are not available in WordNet but it relates to the WordNet concepts.

In addition, [28] presented a method to extend the Portuguese and Galician WordNet's (PULO and Galnet) by exploring a terminological database (Termoteca) by adding new definitions, word as concept, Synset glosses and variants. The proposed method aligned the concepts of WordNet and Termoteca. Taking the WordNet synsets and the concept of terminological records along with the knowledge domains and morphological category is used to reduce the occurrence of homography and polysemy in WordNet domains Hierarchy and Termoteca field domains. The works in [29] used query classification framework that performs the automatic identification and classification of user queries in a specified domain by analyzing the syntactical pattern.

Information retrieval model has also been used in [30] to give answer to the users' query based on knowledge base. This model incorporates domain specific approach by using knowledge-based model and converting the knowledge into quantitative values, which represent the relationship of terms. The knowledge level has been reduced by representing the knowledge into statistical model like Vector space model. The proposed model also defined the knowledge reduction method by representing the knowledge sing semantic network which makes the knowledge representation model more flexible.

Furthermore, a multi-purpose framework was presented in [31] that exploits the generic knowledge base of WordNet to address the issue of "semantic heterogeneity" between ontologies. The multi-purpose model was presented for: (1) In the first step the incorrectly semantic relation between the concepts of ontology is discovered and corrected in a specific domain. (2) Compute the semantic relations to merge with domain-specific ontologies. (3) Acquisition of statistical information to tackle the WordNet missing concept issue. (4) Incorporating the missing concepts into WordNet. The work in[32] implemented multi-classifier with improved architecture by training data and word embeddings in the context of constructing financial guidance natural language understanding (NLU) system. The works in [33] extended WordNet for the acquisition domain specific knowledge from text.

In other work, pre-trained domain specific model for patent domain was presented in [34] using word embeddings. The model was trained on more than 5 million patents to evaluate the classification process. The proposed model used deep learning approach for automatic classification on word embedding.

To acquire the knowledge in new domains for conceptual sentence analysis, Kenmore framework based on domain specific acquisition is presented in [35]that exploits on-line corpus with the minimal human intervention. The result was presented by using the proposed framework in two real-world domains to address a range of following sub problems: (1) POS tagging, concept tagging and semantic feature tagging for all words in corpus. (2) To obtain heuristics for disambiguation of POS and Semantic feature and concept activation. (3) To discover the antecedents of relevant pronouns.

The work in [36] described the innovative approach of enriching WordNet (sloWNet) by expending its domain exploiting multiple resources. The proposed method used Slovene corpus of informatics domain to produce terminology and as a bilingual resource the parallel English-Sloven corpus and online dictionary for mapping the terms to sloWNet. The English term is identified through WordNet 2.1 and then translation is performed by using bilingual lexicon into Slovene. Next, a hybrid approach is used for the extraction from Slovene domain-specific corpus and finally match with existing WordNet synsets.

Similarly,[37] presented a framework that extracts concept from biomedical documents. The proposed framework aims to improve document retrieval by transforming queries and documents from term to concept space and estimate the relevance model by analyzing the concepts semantically. Three main advantages of this approach include: (1) the words and concept have strong dependency (2) reduce the dimensionality of space, unifies the concept and reliable estimates of the relevance. (3) For domain resources, it performs the semantic analysis for accurate distribution of relevance.

Another work in [38] employed ontology learning framework based on WordNet and web documents. This framework extracts relevant words from these resources and demonstrated for biological domain. Two techniques have been explored in this work by expanding the vocabulary. Lexical expansion of WordNet and text mining to extract the vocabulary for ontology and similarity computation is performed to extract the most relevant since fin given concept word.

Different approaches such as knowledge-based model, ontology base model, information retrieval, learning frameworks, domain specific model and hybrid model have been reviewed. However, there is a need to explore the applicability of rule-based model to acquire domain specific features for determination of semantic similarity in the judicial domain.

3. Methodology

This section presents the proposed method of acquiring domain specific features using rule-based approach. This section is comprised of following components: (i) acquisition of domain specific features, (ii) detail process flow of the proposed model, and (iii) Rules of acquisition.

3.1 Acquisition of domain specific features

One important issue in feature selection process is to acquire the domain specific features. In this work, a method of acquiring a domain specific feature is proposed and implemented, which is based on domain independent lexicon WordNet 3.1 and domain specific Black Law dictionary. To obtain the domain specific features from the prepared input query, a rule-based method is designed and developed. The set of rules are applied to the content words i.e., Noun, Verb, Adverb and Adjective in the prepared input query to get the updated domain specific features set. The overall process of acquisition of domain specific features is presented in figure 1)

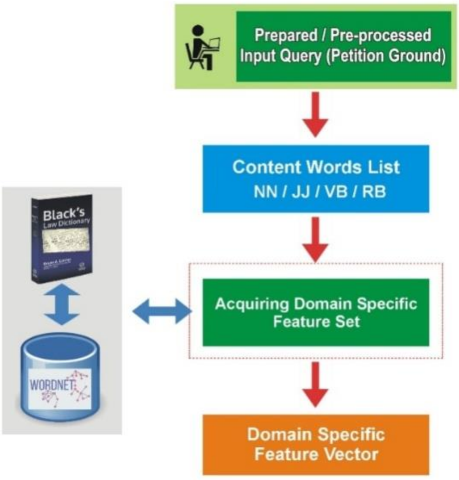

Figure 1. overall process of acquiring domain specific features

3.2 Detail process flow of the proposed model

Many terms in specialized domain do not exist in general purpose lexicon. The proposed model is designed for specific domain i.e., judicial domain. The proposed process flow of acquiring domains specific features is illustrated in figure 2.

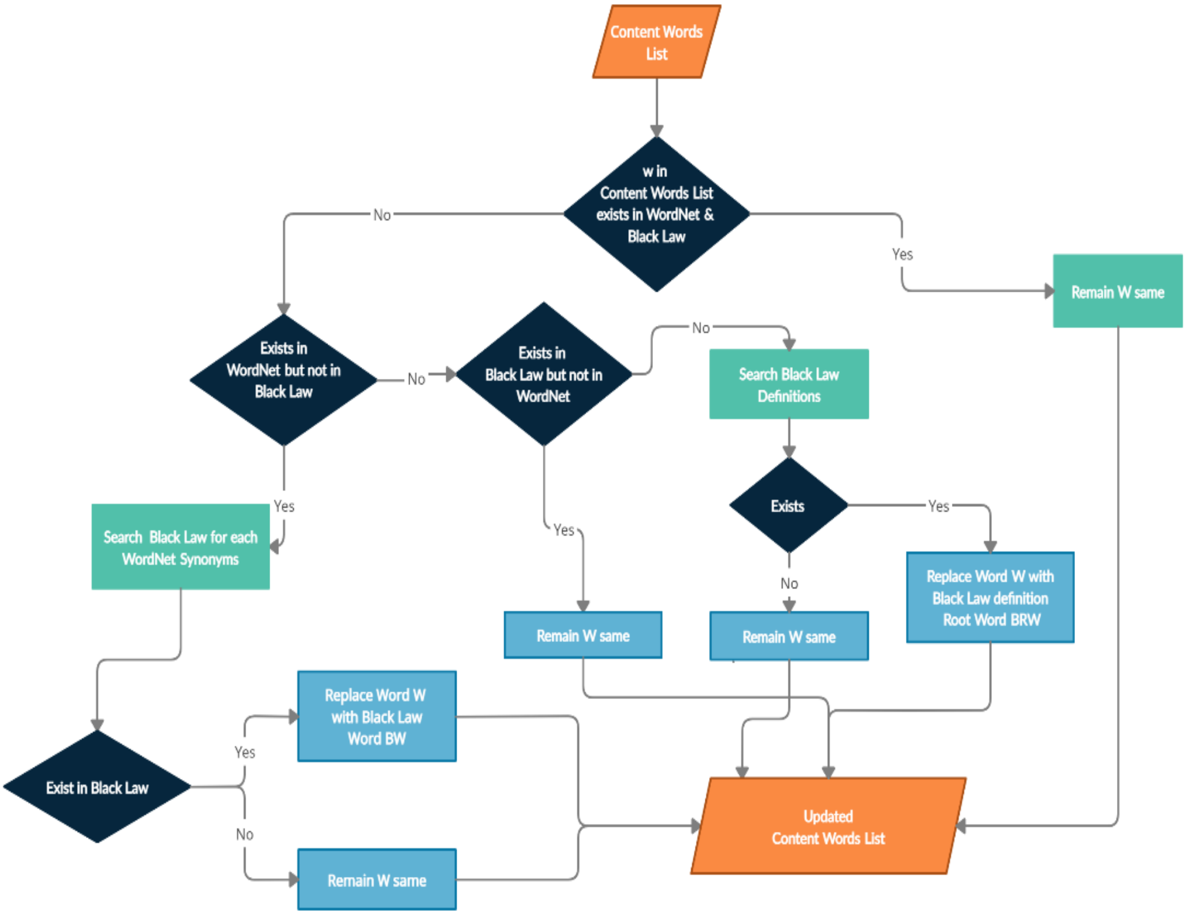

Figure 2. detail process flow of the proposed model

3.3 *Rules of Acquisition*

This section gives a detail description of the set of rules applied in the proposed approach to the content words list to acquire the domain specific feature list. Following are the rules.

ule 1: **IF** word w_i of content words list exists in both WordNet and Black Law **THEN** word w_i remains same and check for next word w_{i+1}.

ule 2: *(a)* **IF** word w_i of content words list exists in WordNet but does not exists in Black Law **THEN** search the word synonyms ws of word w_i in Black Law:

(b) IF found in Black Law **THEN** update word w_i with Black Law Word bw_i and update the content words list.

(c) IF not found **THEN** word w_i remains same and check for next word w_{i+1}.

Rule 3: **IF** word w_i of content words list exists in Black Law but not in WordNet **THEN** word w_i remains same and check for next word w_{i+1}.

Rule 4: *(a)* **IF** word w_i of content words list does not exist in both WordNet and Black Law **THEN** search the word w_i in Black Law definitions:

(b) IF found in Black Law definitions with inverted comma **THEN** update word w_i with Black Law root word brw_i and update the content words list.

(c) IF not found **THEN** word w_i remains same and check for next word w_{i+1}.

The algorithm of the proposed Rule-base method is presented in below Algorithm 1.

Algorithm 1: Acquiring Domain Specific Features

Input: WordNet,BlackLaw,Content Words List CWL[w$_i$, $w_{i+1}w_{i+n}$]
Output: Updated Content Words List CWL[w$_i$, $w_{i+1}w_{i+n}$]
for *each*, $wi \in CWL$ **do**
  // Search in WordNet and BlackLaw
  if $wi \in WordNet$ $and\ wi \in BlackLaw$ **then**
    $wi = wi$
  else if $wi \in WordNet$ $but\ wi \notin BlackLaw$ **then**
    Search WordNet Synonyms ws of wi in $BlackLaw$
    if $ws \in BlackLaw$ **then**
      $wi \leftarrow bw$
    else
      $wi = wi$
  else if $wi \in BlackLaw$ $and\ wi \notin WordNet$ **then**
    $wi = wi$
  else if $wi \notin WordNet$ $and\ wi \notin BlackLaw$ **then**
    Search wi in Black Law Definitions bdw
    if $wi \in bdw$ **then**
      Update wi with corresponding Black Law root word *brw* of *bdw*
      $wi \leftarrow brw$
  else
    $wi = wi$

4. Results and Discussion

This section presents the experimental results of proposed rule-base model obtained while addressing each rule. Table 1 presents the sample input query of the judicial domain. After pre-processing we got the pre-processed Content Words List.

Table 1. Sample Input Query

Category	Year	Query
Criminal	2017	That the petitioner is sui Juris and never ever being engaged in any criminal proceedings till. the present time and by profession, the petitioner is a Medical Technician and performed his duty at Rescue1122 at City XYZ with honesty and regularly and the FIR is totally illegal, ineffective, also bear no authenticity regarding the factual and circumstances of the case, hence liable to be quashed in the light of the available documents circumstances.. That the entire documents reveals that the FIR mentioned above is lodged with mala fide intention and only for dragging the petitioner in the illegal case. That the story in FIR is frivolous, false so the FIR in question is liable to be quashed in the interest of justice. That the instant FIR cannot be sustained in the eye of law as no offence has been committed by the present petitioner and the FIR is tainted with malice I nothing but abuse of law and further a futile exercise. That there is no other efficacious alter adequate remedy except to invoke extra ordinary constitutional jurisdiction ·of this Hon'able court. Any other grounds may also be taken during the course of arguments with the permission of this Hon'able Court. It is therefore, most humbly prayed that on acceptance of the instant quashment petition, the FIR No. XXXX Dated: 21/11/2016 US 70, 419, 420, 468, 471, 109 PPC Police Station City XYZ may kindly be quashed.

Applying rules to the content words list of the sample input query shown in Table 1, Table 2, shows the result of obtaining domain specific features along with rules applied.

Table 2. Identified sample domain specific features

Feature	**Applied Rule No**
Sui Juris	Rule No.3
Mala fide	Rule No.4
Efficacious	Rule No.2
Quashed	Rule No.1

The proposed method is evaluated on a total of 600 input queries including 150 input queries from Criminal, Civil and Services Categories each and 50 input queries from Revenue, Constitutional and Corporate each. Table 3 and Figure 3 summarizes the results obtained during the process of acquiring domain specific features by applying the rules described in section Rules of Acquisition. The table 3 depicts the total number of input query in each category, the total number of content words and the domain specific features acquired using the define rules.

Table 3. Rule-wise results of acquiring domain specific features

Category	No of Input Query	No. of Content Words	Applied Rules Numbers							
			1	2 (a)	2 (b)	2 (c)	3	4 (a)	4(b)	4 (c)
Criminal	150	24337	22049	593	470	123	1109	586	491	95
Civil	150	21389	19828	310	266	54	832	419	357	62
Revenue	50	8135	7581	147	135	12	320	86	51	35
Constitutional	50	6778	6032	215	200	15	435	96	37	59
Service	150	19701	18086	302	217	85	1037	276	198	78
Corporate	50	5945	5173	111	75	36	460	202	140	62
Total	600	86285	78749	1678	1363	325	4193	1665	1274	391

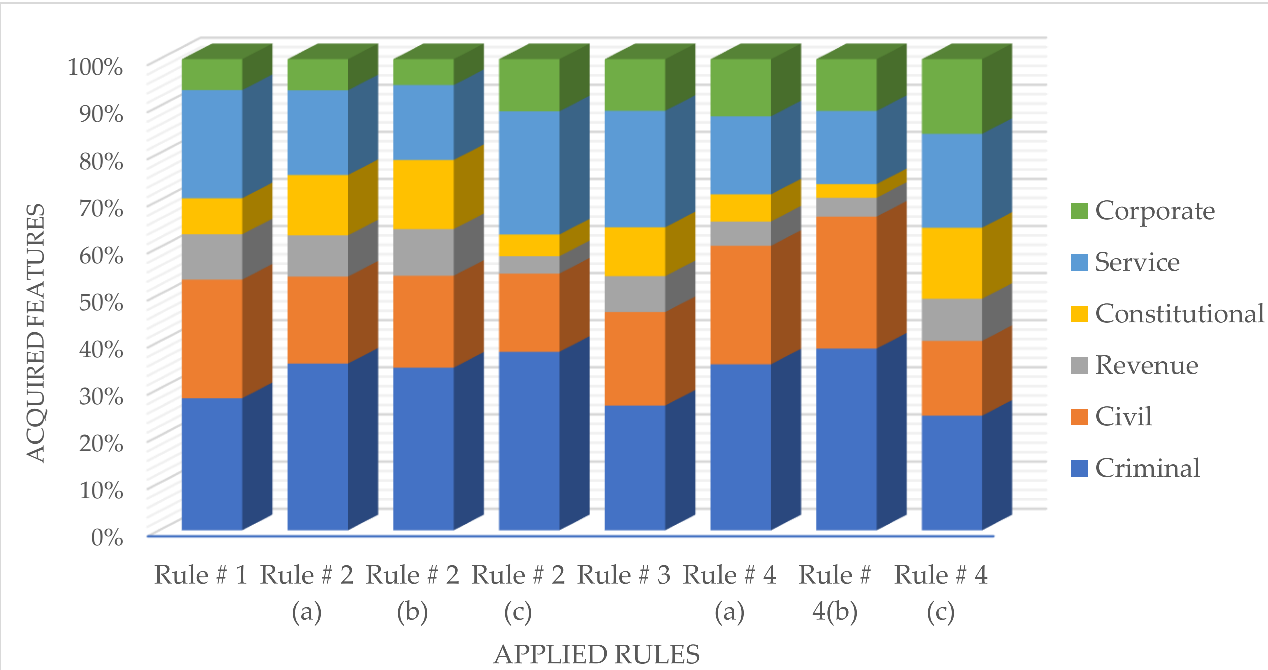

Figure 3 Rule-wise Domain specific features

Table 4 and Figure 4 summarizes the results in terms of the total number of content words in each category and after applying rules to obtain the domain specific features. The last column shows the percentage of the acquired domain specific features in each category. It means that by using the domain specific lexicon i.e., Black Law Dictionary with general purpose lexicon i.e., WordNet, the process of features selection in this domain is enhanced by an average of 9.28%.

Table 4. overall acquired domain specific features

Category	No of Input Query	No. of Content Words	Acquired Domain Specific Features	%age
Criminal	150	24337	2288	9.40
Civil	150	21389	1561	7.29
Revenue	50	8135	553	6.79
Constitutional	50	6778	746	11.00
Service	150	19701	1615	8.19
Corporate	50	5945	773	13.00

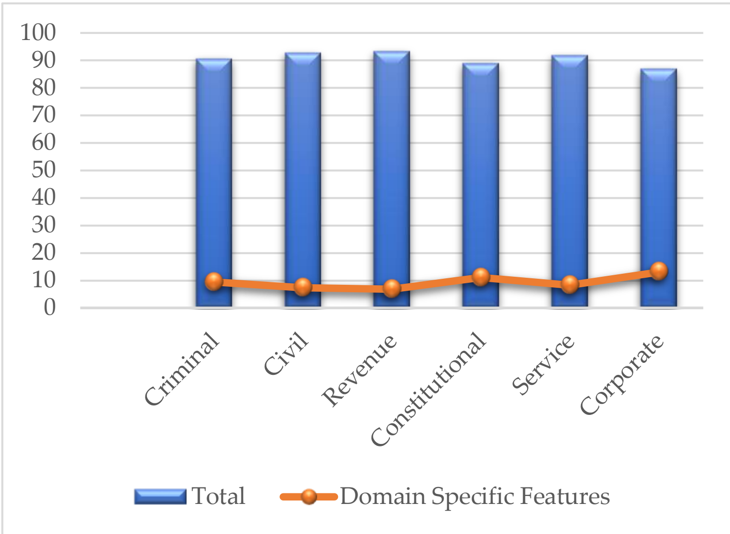

Figure 4. overall acquired domain specific features

5.Performance of the proposed method

The proposed rule-based model, when applied to the preprocessed input query, the performance results in terms of accuracy, precision, recall and F1 measure are shown in table 5.

Table 5. performance of the proposed method

Accuracy	Precision	Recall	F1 Measure
90.2%	0.92	0.91	0.90

6. Discussion

The experiment proved that acquiring domain specific features have significant impact on the results. The proposed rule-based model extracted the features in the judicial domain which will increase in accuracy in determination of semantic similarity. By utilizing the proposed rule-based model, on average 9.28% features are updated in 6 different categories. A more domain specified feature set will have more accuracy in determination of semantic similarity.

7. Conclusion and Future Work

In this research work, a rule-based model was applied for the acquisition of domain specific feature from the content words list of the input query and updates the feature list. The proposed rule-based model performed the following tasks: (i) content words list extraction from preprocessed input query (ii) acquisition of domain specific feature extraction utilizing WordNet and Black Law dictionary (iii) updating the feature list and preparation of input vector for LSTM model.

In the proposed rule-based model, Black Law dictionary was combined with WordNet lexical resource to handle the issue of different context and meaning of same feature in Judicial domain and in the general-purpose domain. From evaluation and experimental results, the proposed rule-based model achieved an accuracy (90.2%), where precision is (0.92), recall is (0.91) and F-measure is (0.90, which is remarkable contribution to the Judicial process and will get good results in real time applications.

7.1 Limitations

The proposed method has some limitation which are described as follows:

- Applying rules to each feature of the content words list reduces the performance of the model.
- Applying Rule No. 4 part (b), if multiple words are found in black law definition with inverted comma, the proposed model selects the first available feature and updatesit in the feature list.
- The proposed study is limited to only English Language and Judicial Domain.

7.2 Future Work

Following are the suggestions for future work:

- Applying rules to only important features can increase the performance.
- Utilizing Rule No.4, feature can be accurately extracted by checking the context of each feature with inverted comma in black law definition and select only the one that has nearest context.
- This work can be probably be enhanced by exploring other legal dictionaries of judicial domain.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All relevant data are within the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Harispe, S.; Ranwez, S.; Janaqi, S.; Montmain, J. Semantic Similarity from Natural Language and Ontology Analysis. *Synthesis Lectures on Human Language Technologies*, **2015**, *8* (1), 1–254. https://doi.org/10.2200/s00639ed1v01y201504hlt027.
2. Maind, A. Measurement of Semantic Similarity Between Words: A Survey. *International Journal of Computer Science, Engineering and Information Technology*, **2012**, *2* (6), 51–60. https://doi.org/10.5121/ijcseit.2012.2605.
3. Mueller, J.; Thyagarajan, A. Siamese Recurrent Architectures for Learning Sentence Similarity. *Proceedings of the AAAI Conference on Artificial Intelligence*, **2016**, *30* (1).
4. Janowicz, K.; Keßler, C.; Schwarz, M.; Wilkes, M.; Panov, I.; Espeter, M.; Bäumer, B. Algorithm, Implementation and Application of the SIM-DL Similarity Server. In *GeoSpatial Semantics*; Springer Berlin Heidelberg: Berlin, Heidelberg; pp 128–145. https://doi.org/10.1007/978-3-540-76876-0_9.
5. Ballatore, A.; Bertolotto, M.; Wilson, D. C. Geographic Knowledge Extraction and Semantic Similarity in OpenStreetMap. *Knowledge and Information Systems*, **2013**, *37* (1), 61–81. https://doi.org/10.1007/s10115-012-0571-0.
6. Budanitsky, A.; Hirst, G. Semantic Distance in Wordnet: An Experimental, Application-Oriented Evaluation of Five Measures. *Workshop on WordNet and Other Lexical Resources*, **2001**.
7. Hasan, A. M.; Mohd Noor, N.; Rassem, T. H.; Mohd Noah, S. A.; Hasan, A. M. A Proposed Method Using the Semantic Similarity of WordNet 3.1 to Handle the Ambiguity to Apply in Social Media Text; 2020; pp 471–483. https://doi.org/10.1007/978-981-15-1465-4_47.
8. Araque, O.; Zhu, G.; Iglesias, C. A. A Semantic Similarity-Based Perspective of Affect Lexicons for Sentiment Analysis. *Knowledge-Based Systems*, **2019**, *165*, 346–359. https://doi.org/10.1016/j.knosys.2018.12.005.
9. Pedersen, T.; Pakhomov, S. V. S.; Patwardhan, S.; Chute, C. G. Measures of Semantic Similarity and Relatedness in the Biomedical Domain. *Journal of Biomedical Informatics*, **2007**, *40* (3), 288–299. https://doi.org/10.1016/j.jbi.2006.06.004.
10. Liu, B.; Zhang, L. A Survey of Opinion Mining and Sentiment Analysis. In *Mining Text Data*; Springer US: Boston, MA, 2012; pp 415–463. https://doi.org/10.1007/978-1-4614-3223-4_13.

11. Turney, P. D.; Pantel, P. From Frequency to Meaning: Vector Space Models of Semantics. *Journal of Artificial Intelligence Research*, **2010**, *37*, 141–188. https://doi.org/10.1613/jair.2934.
12. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. **2013**.
13. Church, K. W.; Hanks, P. Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics*, **1990**, *16* (1), 22–29.
14. Gligorov, R.; ten Kate, W.; Aleksovski, Z.; van Harmelen, F. Using Google Distance to Weight Approximate Ontology Matches. In *Proceedings of the 16th International Conference on World Wide Web*; WWW '07; Association for Computing Machinery: New York, NY, USA, 2007; pp 767–776. https://doi.org/10.1145/1242572.1242676.
15. Pennington, J.; Socher, R.; Manning, C. Glove: Global Vectors for Word Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2014; pp 1532–1543. https://doi.org/10.3115/v1/D14-1162.
16. Levy, O.; Goldberg, Y.; Dagan, I. Improving Distributional Similarity with Lessons Learned from Word Embeddings. *Transactions of the Association for Computational Linguistics*, **2015**, *3* (0), 211–225.
17. Sánchez, D.; Batet, M.; Isern, D.; Valls, A. Ontology-Based Semantic Similarity: A New Feature-Based Approach. *Expert Systems with Applications*, **2012**, *39* (9), 7718–7728. https://doi.org/10.1016/j.eswa.2012.01.082.
18. Khan, M. A.; Khan, M. A.; Ahmed, F.; Mittal, M.; Goyal, L. M.; Jude Hemanth, D.; Satapathy, S. C. Gastrointestinal Diseases Segmentation and Classification Based on Duo-Deep Architectures. *Pattern Recognition Letters*, **2020**, *131*, 193–204. https://doi.org/10.1016/j.patrec.2019.12.024.
19. Sharif, M. I.; Li, J. P.; Khan, M. A.; Saleem, M. A. Active Deep Neural Network Features Selection for Segmentation and Recognition of Brain Tumors Using MRI Images. *Pattern Recognition Letters*, **2020**, *129*, 181–189. https://doi.org/10.1016/j.patrec.2019.11.019.
20. Hussain, N.; Khan, M. A.; Sharif, M.; Khan, S. A.; Albesher, A. A.; Saba, T.; Armaghan, A. A Deep Neural Network and Classical Features Based Scheme for Objects Recognition: An Application for Machine Inspection. *Multimedia Tools and Applications*, **2020**. https://doi.org/10.1007/s11042-020-08852-3.
21. Rashid, M.; Khan, M. A.; Sharif, M.; Raza, M.; Sarfraz, M. M.; Afza, F. Object Detection and Classification: A Joint Selection and Fusion Strategy of Deep Convolutional Neural Network and SIFT Point Features. *Multimedia Tools and Applications*, **2019**, *78* (12), 15751–15777. https://doi.org/10.1007/s11042-018-7031-0.
22. Khan, M. A.; Javed, K.; Khan, S. A.; Saba, T.; Habib, U.; Khan, J. A.; Abbasi, A. A. Human Action Recognition Using Fusion of Multiview and Deep Features: An Application to Video Surveillance. *Multimedia Tools and Applications*, **2020**. https://doi.org/10.1007/s11042-020-08806-9.
23. Khan, M. A.; Sharif, M.; Akram, T.; Raza, M.; Saba, T.; Rehman, A. Hand-Crafted and Deep Convolutional Neural Network Features Fusion and Selection Strategy: An Application to Intelligent Human Action Recognition. *Applied Soft Computing*, **2020**, *87*, 105986. https://doi.org/10.1016/j.asoc.2019.105986.
24. Batool, F. E.; Attique, M.; Sharif, M.; Javed, K.; Nazir, M.; Abbasi, A. A.; Iqbal, Z.; Riaz, N. Offline Signature Verification System: A Novel Technique of Fusion of GLCM and Geometric Features Using SVM. *Multimedia Tools and Applications*, **2020**. https://doi.org/10.1007/s11042-020-08851-4.
25. Mansoor, M.; Rehman, Z. ur; Shaheen, M.; Khan, M. A.; Habib, M. Deep Learning Based Semantic Similarity Detection Using Text Data. *Information Technology And Control*, **2020**, *49* (4), 495–510. https://doi.org/10.5755/j01.itc.49.4.27118.
26. Verma, D.; Muralikrishna, S. N. Semantic Similarity between Short Paragraphs Using Deep Learning. In *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*; IEEE, 2020; pp 1–5. https://doi.org/10.1109/CONECCT50063.2020.9198445.
27. Moldovan, D. I.; Girju, R. Domain-Specific Knowledge Acquisition and Classification Using Wordnet. In *Proceedings of the Thirteenth International Florida Artificial Intelligence Research Society Conference, May 22-24, 2000, Orlando, Florida, USA*; Etheredge, J. N., Manaris, B. Z., Eds.; AAAI Press, 2000; pp 224–228.
28. Simões, A.; Guinovart, X. G. Acquiring Domain-Specific Knowledge for Wordnet from a Terminological Database. In *OpenAccess Series in Informatics*; Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing, 2019; Vol. 74. https://doi.org/10.4230/OASIcs.SLATE.2019.6.
29. Mohasseb, A.; Bader-El-Den, M.; Liu, H.; Cocea, M. Domain Specific Syntax Based Approach for Text Classification in Machine Learning Context. In *2017 International Conference on Machine Learning and Cybernetics (ICMLC)*; IEEE, 2017; pp 658–663. https://doi.org/10.1109/ICMLC.2017.8108983.
30. Yoon, C.; Dankel, D. D. Domain Specific Knowledge-Based Information Retrieval Model Using Knowledge Reduction. In *FLAIRS Conference*; 2005.
31. Mohammed, M.; Mohammed, B. On the Merging of Domain-Specific Heterogeneous Ontologies Using Wordnet and Web Pattern-Based Queries. *Journal of Information & Knowledge Management*, **2011**, *10* (01), 23–36. https://doi.org/10.1142/S0219649211002808.
32. Jenset, G. B.; McGillivray, B. Enhancing Domain-Specific Supervised Natural Language Intent Classification with a Top-down Selective Ensemble Model. *Machine Learning and Knowledge Extraction*, **2019**, *1* (2), 630–640. https://doi.org/10.3390/make1020037.

33. Moldovan, D.; Girju, R.; Rus, V. Domain-Specific Knowledge Acquisition from Text. In *Proceedings of the sixth conference on Applied natural language processing -*; Association for Computational Linguistics: Morristown, NJ, USA, 2000; pp 268–275. https://doi.org/10.3115/974147.974184.
34. Risch, J.; Krestel, R. Domain-Specific Word Embeddings for Patent Classification. *Data Technologies and Applications*, **2019**, *53* (1), 108–122. https://doi.org/10.1108/DTA-01-2019-0002.
35. Cardie. Domain-Specific Knowledge Acquisition for Conceptual Sentence Analysis. *Computer Science Department Faculty Publication Series*, **1994**.
36. Vintar, Š.; Fišer, D. Enriching Slovene Wordnet with Domain-Specific Terms. Annotation, exploitation and evaluation of parallel corpora, Language Science Press, Berlin, pp. 35–53. February 2017. https://doi.org/10.5281/zenodo.283489.
37. Wang, C. *Knowledge Extraction and Retrieval for Domain-Specific Documents*; 2015.
38. Luong, H.; Gauch, S.; Wang, Q. Ontology Learning Using Wordnet Lexical Expansion and Text Mining. In *Theory and Applications for Advanced Text Mining*; Sakurai, S., Ed.; IntechOpen: Rijeka, 2012. https://doi.org/10.5772/51141.