Procesamiento del Lenguaje Natural, Revista nº 51, septiembre de 2013, pp 49-56 recibido 30-04-2013 revisado 18-06-2013 aceptado 21-06-2013

Extracting Drug-Drug Interaction from Text Using Negation Features

Estudio del efecto de la negación en la detección de interacciones entre fármacos

Behrouz Bokharaeian, Alberto Díaz
NIL Group
Universidad Complutense de Madrid
Madrid, Spain
behroubo@ucm.es, albertodiaz@fdi.ucm.es

Miguel Ballesteros
Natural Language Processing Group
Universitat Pompeu Fabra
Barcelona, Spain
miguel.ballesteros@upf.edu

Resumen: La extracción de relaciones entre entidades es una tarea muy importante dentro del procesamiento de textos biomédicos. Se han desarrollado muchos algoritmos para este propósito aunque sólo unos pocos han estudiado el tema de las interacciones entre fármacos. En este trabajo se ha estudiado el efecto de la negación para esta tarea. En primer lugar, se describe cómo se ha extendido el corpus DrugDDI con anotaciones sobre negaciones y, en segundo lugar, se muestran una serie de experimentos en los que se muestra que tener en cuenta el efecto de la negación puede mejorar la detección de interacciones entre fármacos cuando se combina con otros métodos de extracción de relaciones.
Palabras clave: Interacciones entre fármacos, negación, funciones kernel, máquinas de vectores de soporte, funciones kernel.

Abstract: Extracting biomedical relations from text is an important task in BioMedical NLP. There are several systems developed for this purpose but the ones on Drug-Drug interactions are still a few. In this paper we want to show the effectiveness of negation features for this task. We firstly describe how we extended the DrugDDI corpus by annotating it with the scope of negation, and secondly we report a set of experiments in which we show that negation features provide benefits for the detection of drug-drug interactions in combination with some simple relation extraction methods.
Keywords: Drug-Drug interaction, Negation, Support vector machines, kernel-based methods

1. Introduction

A drug-drug interaction (DDI) occurs when one drug affects the level or activity of another drug, this may happen, for instance, in the case of drug concentrations. This interaction can result on decreasing its effectiveness or possibly altering its side effects that may even the cause of health problems to patients (Stockley, 2007).

There is a great amount of DDI databases and this is why health care experts have difficulties to be kept up-to-date of everything published on drug-drug interactions. This fact means that the development of tools for automatically extracting DDIs from biomedical resources is essential for improving and updating the drug knowledge and databases.

There are also many systems on the extraction of biomedical relations from text; however the research on studying the effect of negation in biomedical relation extraction is still limited. On the other hand, negation is very common in clinical texts and it is one of the main causes of making errors in automated indexing systems (Chapman et al., 2001); the medical personnel is mostly trained to include negations in their reports. Particularly when we are detecting the interaction between drugs, the presence of negations can produce false positives classifications, for instance, the sentence *Coadministration of multiple doses of 10 mg of lenalidomide had no effect on the single dose pharmacokinetics of R- and S- warfarin* a DDI between *lenalidomide* and *warfarin* could be detected as a practicable fact if negation is not taken into account. We therefore believe that detecting the words that

ISSN 1135-5948

are affected by negations may be an essential part in most biomedical text mining tasks that try to obtain automatically the accurate knowledge from textual data.

In order to avoid errors derived of using automatic negation detection algorithms such as NegEx (Amini et al., 2011), we annotated a DDI corpus - previously developed- with the scope of negations. The corpus is called DrugDDI corpus (Segura-Bedmar et al., 2011b), and it was developed for the Workshop on Drug-Drug Interaction Extraction (Segura-Bedmar et al., 2011a) that took place in 2011 in Huelva, Spain. The DrugDDI corpus contains 579 documents extracted from the DrugBank database. We analyzed the corpus and we annotated the sentences within with the scope of negation in order to find the effect of negation features in the detection of DDIs. We annotated it using the same guidelines of the BioScope corpus (Vincze et al., 2008), that is, we annotated cues and scopes affected by negation statements into sentences in a linear format.

For detecting the DDIs we used a fast version of a support vector machine (henceforth, SVM) classifier with a linear kernel based on a bag of words (henceforth, BOW) representation obtained from the extracted features. We carried out some experiments with different kernels (global context, subtree, shortest path), with and without negation information. The results presented in this paper show that negation features can improve the performance of relation extraction methods.

The rest of the paper is structured as follows. In Section 2 we discuss previous related work about biomedical relation extraction and relevant information about the DrugDDI corpus. In Section 3 we explain how we annotated the corpus with the scope of negation. In Section 4 we explain how we used the obtained information from negation tags to improve the DDI detection task. In Section 5 we discuss the results obtained. Finally, in Section 6, we show our conclusions and suggestions for future work.

2. Related work

In this Section we describe the DrugDDI corpus and we present some related work on kernel-based relation extraction.

2.1. DrugDDI corpus

There are some annotated corpora that were developed with the intention of studying biomedical relation extraction, such as, Aimed (Bunesu et al., 2005), LLL (Nedellec et al., 2005), BioCreAtIvE-PPI (Krallinger et al., 2008) on protein-protein interactions (PPI) and DrugDDI (Segura-Bedmar et al., 2011b), on drug-drug interactions. In particular, the DrugDDI corpus is the first annotated corpus on the phenomenon of interactions among drugs and it is the one that we used for our experiments. It was designed with the intention of encouraging the NLP community to conduct further research on this type of interactions. The DrugBank database (Wishart et al., 2008) was used as source to develop this corpus. This database contains unstructured textual information on drugs and their interactions.

The DrugDDI corpus is available in two different formats: (i) the first one contains the information provided by MMTX (Aronson, 2001) and the unified format adapted from PPI corpora format proposed in (Pyysalo et al., 2008). The unified XML format (see Figure 1) does not contain any linguistic information; it only provides the plain text sentences, their drugs and their interactions. Each entity (drug) includes reference (origId) to the sentence identifier in the MMTX format corpus. For each sentence contained in the unified format, the annotations correspond to all the drugs entities and the possible DDI candidate pair that represents the interaction. Each DDI candidate pair is represented as a *pair ID* node in which the identifiers of the interacting drugs are registered on its *e1* and *e2* attributes. If the pair is a DDI, the *interaction* attribute is set to *true*, otherwise this attribute is set to *false*. Table 1 shows related statistics of the DrugDDI corpus (Segura-Bedmar et al., 2011b).

2.2. Biomedical Relation Extraction

Nowadays, there are many systems developed for extracting biomedical relations from text that can be categorized in (i) feature based and (ii) kernel-based approaches. Feature-based approaches transform the context of entities into a set of features; this set is used to train a data-driven algorithm. On the other hand, kernel-based approaches are

```
−<sentence id="DrugDDI.d346.s0" origId="s0" text="Uricosuric Agents: Aspirin may decrease the effects of probenecid,
 sulfinpyrazone, and phenylbutazone.">
   <entity id="DrugDDI.d346.s0.e0" origId="s0.p0" charOffset="0-17" type="drug" text="Uricosuric Agents"/>
   <entity id="DrugDDI.d346.s0.e1" origId="s0.p2" charOffset="19-26" type="drug" text="Aspirin"/>
   <entity id="DrugDDI.d346.s0.e2" origId="s0.p6" charOffset="55-65" type="drug" text="probenecid"/>
   <pair id="DrugDDI.d346.s0.p0" e1="DrugDDI.d346.s0.e0" e2="DrugDDI.d346.s0.e1" interaction="false"/>
   <pair id="DrugDDI.d346.s0.p1" e1="DrugDDI.d346.s0.e0" e2="DrugDDI.d346.s0.e2" interaction="false"/>
   <pair id="DrugDDI.d346.s0.p4" e1="DrugDDI.d346.s0.e1" e2="DrugDDI.d346.s0.e2" interaction="true"/>
</sentence>
```

Figure 1: The unified XML format in the DrugDDI corpus.

	No.
Documents	579
Sentences	5,806
Drugs	8,260
Sentences with at least two drug	3,775
Sentences with at least one DDI	2,044
Sentences with no DDI	3,762
Candidate drug pairs	30,757
Positive interactions	3,160
Negative interactions	27,597

Table 1: Basic statistics for the DrugDDI corpus.

based on similarity functions. This idea provides the option of checking structured representations, such as parse trees and computing the similarity between different representations directly. Combing kernel based and feature based approaches were investigated by Thomas et al. (2011), they developed a voting system (based on majority) that benefits from the outcomes of several methods.

So far, the sequence and tree kernels are the ones that have shown a superior performance for the detection of biomedical relations from text (Bunesu et al., 2005). In particular, global context kernel, subtree and shortest path kernels are three important kernel methods that were applied successfully for biomedical relation extraction task. For instance, Giuliano et al. (2005) applied by considering three different patterns and they calculated the similarity between two sentences by computing common n-grams of two different patterns.

The shortest path kernel (Bunescu y Mooney, 2005) uses the shortest path between two entities (or drugs) in a phrase structure tree. The subtree kernel (Moschitti, 2006) counted the number of common subtrees in whole parse trees by comparing two different sentences. Moreover, a comparative survey about different kernel-based approaches and their performances can be found in (Frunza y Inkpen, 2010).

More recent research on tree kernels were carried out by Guodong et al. (2010). They introduced a "context-sensitive" convolution tree kernel, which specifies both "context-free" and "context-sensitive" sub-trees by traversing the paths of their ancestor nodes as their contexts to capture structural information in the tree structure. Another motivating work was reported by Chen et al. (2011), that presented a proteing-protein interaction pair extractor, it consists on a SVM classifier that exploits a linear kernel with a complete set of features.

Finally, Simões et al. (2013) introduced an approach for Relation Extraction (RE) based on labeled graph kernels, they proposed an implementation of a random walk kernel (Neuhaus y Bunke, 2006) that mainly explores two characteristics: (i) the words between the candidate entities and (ii) the combined information from different sources.

3. Annotating the DrugDDI corpus with negations

We followed the Bioscope guidelines in order to annotate the corpus (Vincze et al., 2008). The main idea is based in the detection of a set of negation cues, like 'no' or 'not'. After this, the scope of the cue is calculated based on its syntactic context. There are several systems that annotate the scope of negation, in our approach we used the one published by Ballesteros et al. (Ballesteros et al., 2012), which is publicly available,[1] rule-based system that works on biomedical literature (Bioscope) and the input is just the sentence without any required annotation, which serves very well for our purposes.

We used as input all the sentences of the DrugDDI corpus, containing 5,806 sentences and 579 files. The output was therefore a set of sentences annotated with the scope of negation. After applying the system, we observed that there were a set of 1,340 sentences containing negations in the corpus,

[1] `http://minerva.fdi.ucm.es:8888/ScopeTagger/`

which conforms 23% of the corpus.

Taking into account that the negation scope detection system is fully automatic, we manually checked the outcome correcting the annotations when needed. In order to do so, we divided the annotated corpus in 3 different sets that were assigned to 3 different evaluators. The evaluators checked all the sentences contained in each set and corrected the sentences that contained annotation errors. After this revision, a different evaluator revised all the annotations produced by the other 3 evaluators. Finally, we got the whole set of 1,340 sentences (correctly) annotated with the scope of negation.

The algorithm produced errors -according to the evaluators- in 350 sentences from the 1,340, including false positives matches (there were 16 cases). Which means that 74% of sentences was annotated correctly in an automatic way, when considering a full scope match. The main errors produced by the algorithm were related with the processing of passive voice sentences, commas, and copulative keywords (and, or). In particular the problem of passive voice sentences was related with the pattern *It + to be + not + past participle*, which seems that it was not captured by the system, at least in all cases. The false positives were related with the cue *failure*, which is not a negation when it is a noun modified by an adjective, for instance, *renal failure* or *heart failure*. In the DrugDDI corpus these words appear always as nouns, and therefore all of the performed annotations were incorrect.

The following paragraph shows some examples and corrections made by the evaluators:

- Scope closed in an incorrect way containing words from two different clauses such as: Example: *It is [{not} clear whether this represents an interaction with TIKOSYN or the presence of more severe structural heart disease in patients on digoxin;]*. The scope should be closed in or.
- Scope closed in an incorrect way in copulative coordinated sentences: Example: *The following medications have been administered in clinical trials with Simulect? with [{no} increase in adverse reactions: ATG/ALG] , azathioprine, corticosteroids, cyclosporine, mycophenolate mofetil, and muromonab-CD3.*
- Scope opened incorrectly in coordinated copulative sentences: Example: *In an in vitro study, cytochrome P450 isozymes 1A2, 2A6, 2C9, 2C19, 2D6, 2E1, [and 3A4 were{not} inhibited by exposure to cevimeline].*
- Some passive voice sentences were not detected. In particular, as it is already mentioned, sentences with the format 'It (this and that) + finite form of to be + not + past participle'. Example: *[Concomitant use of bromocriptine mesylate with other ergot alkaloids is{not} recommended].*

We also carried out some analysis concerning the number of different cues in the corpus and the number of different errors observed. Table 2 shows that *not* and *no* are by far the most frequent cues in the corpus. It can be observed that the most problematic cue is *neither ... nor*.

Cue	No.	MODFs	Rate
Not	855	266	31.1%
No	439	58	13.2%
without	47	8	17.0%
Neither ... nor	14	12	85.7%
Absence	10	5	50%
Lack	8	1	12.5%
cannot	7	4	57.1%

Table 2: Statistics of negations cues in the corpus and modifications for each cue in the manual checking process.

We finally explored the sentences that are not automatically annotated but they indeed show a negative statement in order to find false negatives. We looked into several negations cues that are not detected by the system such as *unaffected, unchanged* or *nonsignificant*. We detected and corrected 75 different sentences that belong to this problem.

Here we show some examples of false negatives:

- *[The pharmacokinetics of naltrexone and its major metabolite 6-beta-naltrexol were {unaffected} following co-administration with Acamprosate].*
- *[Mean T max and mean plasma elimination half-life of albendazole sulfoxide were {unchanged}].*

- *Monoamine Oxidase Inhibition: Linezolid is a reversible, [{nonselective} inhibitor of monoamine oxidase].*

Therefore, the corpus finally contains 1,399 sentences annotated with the scope of negation, of which 932 correspond to sentences in which there are at least two drugs mentioned. It is worth mentioning that there are 1,731 sentences with 2 or more drug mentions but no DDI, and 2,044 with 2 or more drugs and at least one interaction.

Finally, the extension of the DrugDDI corpus consists of adding a new tag in the annotation of each sentence with the scope of negations. Figure 3 shows an example. The produced corpus is available for public use.[2]

4. DDI detection

In this Section, we explain in detail the experiments we carried out by using negation features. First, we illustrate in detail the methods we used without negation features, and then we present our proposed combined negation method, see figure 4. All the experiments were carried out by using the Stanford parser[3] for tokenization and constituent parsing (Cer et al., 2010), and the SVMs provided by Weka as training engine.

4.1. DDI detection without negation features

The DDI extraction method consists of four different processes: (1) initial prepossessing, (2) feature extraction, (3) Bag of Words computation and (4) classification. The preprocessing step (1) consists of removing some stop words and tokens, for instance removing question marks at the beginning of the sentence. We also carried out a normalization task for some tokens due to the usage of different encoding and processing methods, mainly HTML tags. In the feature extraction step (2), we extracted three different feature sets corresponding to different used relation extraction methods. The feature extraction step for global context kernel consists of extracting *fore-between, between*, and *between-after* tokens that we mentioned in Section 2. The feature extraction step for shortest path kernel method included constituent parsing of the sentence and then extracting shortest path between two drugs in the generated parse tree. And for the subtree kernel we also extracted all subtrees from the mentioned constituent parse tree. After extracting features, we applied the BOW method (3) to generate new feature sets that the SVM classifier uses. The aim of this step is producing a new representation of the instances which is used in the classification step. And finally in the classification step (4), we applied the Weka SVM classifier (Platt, 1998) (SMO), with a linear composition of features produced by the BOW method to detect the interactions among drugs. The Inner product of new features was used as kernel function between two new representations.

4.2. DDI detection with negation features

In this section, we explain our proposed method that merges negation features with the features mentioned in Section 4.1. We divided the corpus in instances affected by negation and instances without negation statements. The last ones were classified in the same way as in Section 4.1, while for the instances with negations we added negation features to the representation. The positive instances were classified in the same way as previous approaches but the sentences containing negations were categorized using negation features in addition to the other previous features. As in previous subsection, the combined method for instances containing negations consists of 4 steps. After a simple preprocessing step we carried out a feature extraction process. In this step, we generated six negation features in addition to three feature sets corresponding to global context kernel (GCK), Shortest Path and Subtree kernel methods. Negation feature consists of tokens inside the negation scope, left side tokens outside of the negation scope and right side tokens, and the negation cue tokens, negation cue, and position of open and closed negation scope. For instance in the sentence shown in Figure 3: tokens inside brackets create middle scope features, right side tokens construct right features and tokens in the left side of the negation scope form left scope features. As in the previous subsection, we used a BOW method to convert negation string features to word features. Finally, the new feature set is used to classify the drug-drug interactions by

[2] `http://nil.fdi.ucm.es/sites/default/files/NegDrugDDI.zip`

[3] `http://nlp.stanford.edu/software/lex-parser.shtml`

```
<sentence origId="s0" id="DrugDDI.d291.s0" text="Zidovudine: There is no significant
pharmacokinetic interaction between ZDV and zalcitabine which has been confirmed
clinically.">
        <entity .... />
        <pair .... />
        <negationtags>Zidovudine: There is <xcope><cue>no</cue> significant pharmacokinetic
            .... clinically</xcope>.</negationtags>
</sentence>
```

Figure 2: A sentence annotated with the Scope of Negation in our version of the DrugDDI corpus.

making use of the Weka SVM.

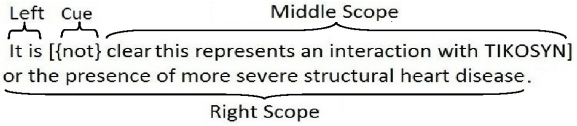

Figure 3: Left, middle and right side scope and negation cue in a negative sentence.

In summation, our approach is a feature based method that uses a bag of word kernel utilizing basic features to compute simple basic kernels and negation features. We applied a fast implementation of the support vector machine provided in Weka, which uses sequential minimal optimization. By carrying out some experiments we also limited the size of the words in each feature bag in the BOW approach to 1000 words per feature class.

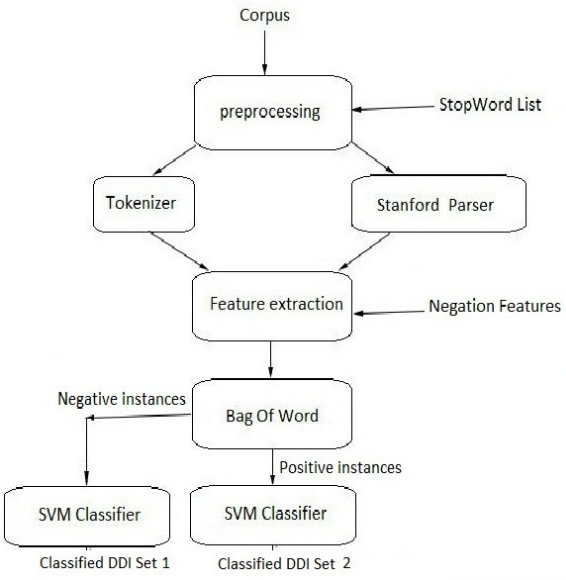

Figure 4: The different processes followed by our proposed approach.

5. *Evaluation*

5.1. Evaluation Setup

In order to demonstrate the improvements provided by using negation features, our experiments consisted of a 10-fold cross validation over the training part of the DrugDDI corpus. Therefore, our results are not directly comparable to the ones provided in the DDI challenge (Segura-Bedmar et al., 2011a). The training DrugDDI corpus contains 437 documents extracted from DrugBank database. It consists of 4267 sentences with average of 9.8 sentences per document and 25,209 instances with 2,402 interactions between different drugs.

Our measurement metrics included true positive, false positive, false negative, total number of positive instances, Precision, Recall and F-1 score.

5.2. Results

Table 3 shows the outcomes of the experiments by computing the metrics and by training over the DDI corpus. The table shows results for Global context kernel (GCK), Subtree and shortest path kernel (SubtreeK) and corresponding combined negation methods (GCKNS = GCK with negation features; SubtreeKNS = Subtree kernel with negation features). The first three rows of the table show the performance of the three basic kernels and the last three ones (with NS postfix) show the outcomes for the combined version that includes negation features. The best result was obtained with GCKNS, and the worst result was obtained by the shortest path tree approach. Moreover, the best improvement was obtained by the GCK approach; it improves 3.8 percentual points of the F score.

As we can see in the table, there is an improved performance when we applied the negation features for classification. This fact

Method	TP	FP	FN	Total	P	R	F1
GCK:	902	1094	1500	2402	0.452	0.376	0.410
SubtreeK:	818	1105	1584	2402	0.425	0.341	0.378
ShortestPathTK:	795	1066	1607	2402	0.427	0.331	0.373
GCKNS:	987	1021	1415	2402	0.492	0.411	0.448
SubtreeKNS:	919	1280	1483	2402	0.418	0.383	0.399
ShortestPathTKNS:	936	1240	1466	2402	0.430	0.390	0.409

Table 3: 10- cross validation results for the methods that do not use negation features and the methods that use negation features.

demonstrates our hypothesis and the emphasizes the purpose of the present work.

6. Conclusions and Future Work

Due to the huge amount of drug related information in bio-medical documents and the importance of detecting dangerous drug-drug interactions in medical treatments, we believe that implementing automatic Drug-Drug interaction extraction methods from text is critical. The DrugDDI corpus is the first annotated corpus for Dug-Drug interaction tasks used in the DDI Extraction 2011 challenge.

In this paper, after reviewing related work on biomedical relation extraction, we first explained the process of annotating the DrugDDI corpus with negation tags; and then we explored the performance of combing negation features with three simple relation extraction methods. Our results show the superior performance of the combined method utilizing negation features over the three basic experimented relation extraction methods.

However, the experiments also show that the application of negation features can indeed improve the relation extraction performance but the obtained improvement clearly depends on the number and rate of positive and negative relations, rate of negative cues in the corpus, and other relation extraction features. It is also true that combining negation features with a huge number of other features may not improve the performance and even may hurt the final result, and this is why we used a limited number of features. It is therefore obvious that corpora having more sentences with negation cues can benefit more from using negation features.

For further work, we plan to use a different type of annotation such as negation events instead of scopes, and also handling hedge cues and speculative statements in conjunction with negations.

References

Amini, I., M. Sanderson, D. Martinez y X. Li. 2011. Search for clinical records: Rmit at trec 2011 medical track. En *Proceedings of the 20th Text REtrieval Conference (TREC 2011)*.

Aronson, A. 2001. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. En *Proceedings of the AMIA Symposium*, páginas 17–27. URL http://metamap.nlm.nih.gov/.

Ballesteros, M., V. Francisco, A. Diaz, J. Herrera y P. Gervas. 2012. Inferring the scope of negation in biomedical documents. En *Proceedings of the 13th International Conference on Intelligent Text Processing and Computational Linguistics (CICLING 2012)*.

Bunescu, R., R. Ge, R. Kate, E. Marcotte, R. Mooney, A. Ramani y Y. Wong. 2005. Comparative experiments on learning information extractors for proteins and their interactions. *Artificial Intelligence in Medicine*, 33(2):139–155.

Bunescu R. y R. Mooney. 2005. A shortest path dependency kernel for relation extraction. En *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, páginas 724–731.

Cer, D., M. de Marneffe, D. Jurafsky y C. D. Manning. 2010. Parsing to Stanford Dependencies: Trade-offs between speed and accuracy. En *Proceedings of the 7th In-*

ternational Conference on Language Resources and Evaluation(LREC 2010) .

Chapman, W., W. Bridewell, P. Hanbury, G. F. Cooper y B. G. Buchanan. 2001. A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries. *Journal of Biomedical Informactics*, 34(5):301-310.

Chen, Y., F. Liu y B. Manderick. 2011. Extract Protein-Protein Interactions From the Literature Using Support Vector Machines with Feature Selection. *Biomedical Engineering, Trends, Researchs and Technologies*.

Frunza, O. y D. Inkpen. 2010. Extraction of disease-treatment semantic relations from biomedical sentences. *Proceedings of the 2010 Workshop on Biomedical Natural Language Processing*, páginas 91–98.

Giuliano, C., A. Lavelli y L. Romano. 2005. Exploiting shallow linguistic information for relation extraction from biomedical literature. En *Proceedings of the Eleventh Conference of the European Chapter of the Association for Computational Linguistics (EACL-2006)*, páginas 5–7.

Guodong, Z., Q. Longhua y F. Jianxi. 2010. Tree kernel-based semantic relation extraction with rich syntactic and semantic information. *International Journal on Information Sciences*, 180(8):1313–1325.

Krallinger, M., A. Valencia y L. Hirschman. 2008. Linking genes to literature: text mining, information extraction, and retrieval applications for biology. *Genome Biology* , 9(Suppl 2):S8.

Moschitti, A. 2006. Making Tree Kernels Practical for Natural Language Learning. En *Proccedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*.

Nedellec, C. 2004. Machine Learning for Information Extraction in Genomics - State of the Art and Perspectives. *Text Mining and its Applications*, Springer Verlag.

Neuhaus, M. y H. Bunke. 2006. A Random Walk Kernel Derived from Graph Edit Distance. *Lecture Notes in Computer Science*, 4109(5):191-199.

Platt, J. 1998. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. *Advances in kernel methods - Support vector learning*.

Pyysalo, S., A. Airola, J. Heimonen, J. Bjorne, F. Ginter y T. Salakoski. 2008. Comparative analysis of five protein-protein interaction corpora. *BMC bioinformatics*, 9(Suppl 3):S6.

Segura-Bedmar, I., P. Martínez y D. Sánchez-Cisneros. 2011. En *Proceedings of the 1st Challenge task on Drug-Drug Interaction Extraction (DDIExtraction 2011)*. CEUR Workshop Proceedings, Vol. 761.

Segura-Bedmar, I., P. Martínez y C. de Pablo Sánchez. 2011. Using a shallow linguistic kernel for drug-drug interaction extraction. *Journal of Biomedical Informatics*, 44(5):789–804.

Simões, G., D. Matos y H. Galhardas. 2013. A Labeled Graph Kernel for Relationship Extraction. *CoRR*, abs/1302.4874.

Stockley, I. H. 2007. *Stockley's Drug Interaction*. Pharmaceutical Press.

Thomas, P., M. Neves, I. Solt, D. Tikk y U. Leser. 2011. Relation extraction for drug-drug interactions using ensemble learning. En *Proceedings of the First Challenge task on Drug-Drug Interaction Extraction (DDIExtraction 2011)*, pp:11–17.

Vincze, V., G. Szarvas, R. Farkas, G. Mora y J. Csirik. 2008. The BioScope corpus: annotation for negation, uncertainty and their scope in biomedical texts. *BMC Bioinformatics*, 9(Suppl 11):S9.

Wishart, D. R., C. Knox , A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, M. Hassanali. 2008. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.*, 36(Suppl 1):D901-D906.