Soft Computing
https://doi.org/10.1007/s00500-018-3650-7

METHODOLOGIES AND APPLICATION

Word embeddings for negation detection in health records written in Spanish

Sara Santiso[1] · Arantza Casillas[1] · Alicia Pérez[1] · Maite Oronoz[1]

Abstract

This work focuses on the creation of a system to detect negated medical entities in electronic health records (EHRs) written in Spanish. The importance of this task rests on the influence that the negation can have in the automatic understanding of information given that it inverts the truth value of a clause. We explore a novel continuous characterization as an alternative to previous negation extraction approaches based on discrete characterizations. The aim is to increase the ability of the characterization to generalize over discrete features. We also included other features that could be useful for the negation detection task. In addition, the negation detection is approached as a named entity recognition task where we want to find only the negated entities. EHRs are represented by the corresponding embeddings. In addition, this approach is compared with a traditional discrete characterization based on words. These representations are employed by a supervised classifier such as conditional random fields to infer the predictive model. The approach is assessed on health records from different hospitals, namely IxaMed-GS and IULA. The best performance is achieved by virtue of the embedding-based characterization, leading to an f-measure of 75.3 and 81.6 for the IxaMed-GS and IULA corpus, respectively. With this work, we prove that the use of embedding-based representations can also be useful for the detection of negated medical entities.

Keywords Negation detection · Word embeddings · Machine learning · Text mining

1 Introduction

The digitalization of health records makes available vast amounts of data valuable for computer-based data exploration. Moreover, natural language processing (NLP) enables automatic extraction of relevant information from electronic health records (EHRs) and opens a way for clinical decision making. Indeed, models inferred from large-scale observations can drive the detection of adverse drug events (Henriksson et al. 2015; Casillas et al. 2016), hospital- acquired infections (Cohen et al. 2006; Jacobson and Dalianis 2016) or diagnosis extraction and ICD encoding (Pérez et al. 2015; Asker et al. 2016).

Massive data analysis shall have a direct impact on the improvement of patients safety through evidence-based health practices, but it shall also bring other benefits, such as the reduction of manual revision of documents in the documentation and surveillance services. The aim of this work is to explore data-driven techniques suitable for the negation detection in medical documents. We focus on Spanish language, while being official in a wide number of countries, it could be benefited from the development of NLP tools for specific domains as it is the case of clinical text mining. The task in this work is to detect negated medical entities, either diseases (including disorders, signs, symptoms and findings), or drugs (substances, brand names and active ingredients). We do not use a general negation detection approach where first the negation cue is identified and then its scope. We want to detect directly negated medical entities.

The negation inverts the truth value of a sentence or clause. Consequently, the identification of negated entities is crucial to make automatic systems extract and interpret the information correctly. In this case, the negation of medical entities such as disorders is essential to assess health studies, for example for documentation and insurance billing. Nevertheless, its identification entails several challenges. The creation of a supervised system for the negation detection needs a big amount of annotated data to represent the information

Communicated by V. Loia.

✉ Sara Santiso
sara.santiso@ehu.eus

[1] IXA Group, University of the Basque Country (UPV-EHU), Manuel Lardizabal 1, 20080 Donostia/San Sebastián, Spain

Published online: 23 November 2018

in a big-dimensional space, but this is not always possible to acquire it. Above all, in a language different to English, which counts on a variety of tools and resources in comparison with others (Dalianis et al. 2014), and in the clinical domain, where documents such as EHRs are subject to confidentiality agreements. For the negation detection in the clinical domain in Spanish, we only found publicly available the corpora created in (Cruz et al. 2017; Marimon et al. 2017). In addition, in EHRs such as those used for this work, the same concept can be represented with different word forms, using or not abbreviations, or with misspelling errors. This causes that with a symbolic representation based on word forms, one concept can have assigned multiple representations. Both the data sparsity and the lexical variability tend to decrease the performance of models inferred from data.

With the motivation of overcoming these issues, we propose to tackle negation detection by means of robust characterizations derived from continuous spaces (word embeddings). Word embeddings allow to represent words in a continuous space of small dimension and help inference algorithms to achieve better performance by grouping semantically related words (Mikolov et al. 2013). Recent approaches based on machine learning have demonstrated the robustness of distributional semantics and continuous representations of words to tackle data sparsity (Gormley et al. 2015; Nguyen and Grishman 2016; He and Lin 2016).

Our main contribution is the comprehensive experimental layout exploring alternative continuous representations for the recognition of negated entities in Spanish EHRs. The novelty stands on the fact that the task is approached as named entity recognition (NER) using conditional random fields (CRFs) (Guo et al. 2014; Segura-Bedmar et al. 2015; Copara et al. 2016). By contrast to mainstream techniques that detect the negation cues and their scopes, we are not specially interested in the recognition of a broad scope but in the marking of the absence of diseases and drugs. That is to say, the system focuses on the clinical entities that are negated and this makes possible to avoid information that is not relevant for specific tasks.

2 Related work

The early works on detection of negated entities were done using rule-based systems. Among them, the NegEx tool (Chapman et al. 2001) employs exact match with respect to a list of medical entities and a list of negation triggers and labels as negated entities those inside a token window near the negation triggers. NegEx was demonstrated successful for documents written in English, and it was used in different languages such as Swedish (Skeppstedt 2011), French (Deléger and Grouin 2012), German (Cotik et al. 2016) and Spanish (Costumero et al. 2014).

Later, machine learning-based systems were proven able to outperform systems based on regular expression. For example, Cruz et al. (2010) focused on the detection of the negation expressions in the BioScope corpus (Vincze et al. 2008) with both C4.5 and Naive Bayes (NB) algorithms. The f-measure achieved by NegEx was outperformed by 20%.

Morante and Daelemans (2009) created a system based on machine learning to detect the scope of the negation in biomedical texts. The detection process was divided into two phases: (i) identification of the negation signals and (ii) determination of the negation scope. For the negation signal identification, the IGTREE algorithm was used. For the negation scope determination, a meta-classifier that used the predictions of three base classifiers was used: (i) memory based, (ii) support vector machines (SVMs) and (iii) conditional random fields (CRFs). The meta-classifier was also the CRFs algorithm. The system was assessed on the BioScope corpus using tenfold cross-validation. For the identification of negation signals, the mean of the f-measures was 98.74. For the determination of the negation scope, the mean of the f-measures was 89.15. There are other works that followed the two-step approach incorporating POS information (Agarwal and Yu 2010) or with different classifiers (Cruz et al. 2012).

The machine learning approaches were also used in other types of medical documents. For example, Goldin and Chapman (2003) used the Naive Bayes (NB) and Decision Tree (DT) algorithms to detect the negation with 'not' in reports from the Medical Archival System (MARS) repository. For admission notes and discharge summaries written in Chinese, Kang et al. (2017) incorporated word embeddings and character embeddings to a CRF algorithm. Under the five-fold cross-validation, the best f-measure was 99% for the cue detection and 94% for the scope detection, concluding that embedding features help to achieve better performances.

Based on the previous works, our intuition is that for domains and languages with few linguistic resources (as it is the case of Spanish compared to English), word embeddings can help to tackle the sparsity of data, and based on that, we decided to compare symbolic and continuous representations. However, we did not find works in Spanish applying machine learning techniques with embeddings for negation detection. Kang et al. (2017) developed this task with documents written in Chinese and trained the embeddings with word2vec (Mikolov et al. 2013), whereas our embeddings were trained with GloVe (Pennington et al. 2014), which offered better results in works such as (Artetxe et al. 2018), and skipNgram (Ling et al. 2015). In addition, there is a subtle difference between the aforementioned works and ours. We do not focus on the negation extraction broadly; we focus specifically on negated clinical entities. With this boundary, we decided not to tackle the problem as a two-step problem but as a NER approach instead, with the particularity of the use of continuous features.

3 Materials and methods

3.1 Classification

For what we learned from the related works (refer to Sect. 2), we decided to explore the CRFs supervised classifier. CRFs involve a probabilistic framework for labeling and segmenting sequential data. CRFs (Lafferty et al. 2001) construct a conditional model $p(Y|X)$ to create a discriminative framework from the jointly distributed variables X and Y, instead of modeling the marginal $p(X)$. X are observation sequences and Y their corresponding label sequences. That is to say, it takes into account the information of the earlier and later tokens to make the predictions.

Sequences of tokens are characterized so as to help conveying as much valuable information as possible to the CRF. The features may represent attributes at different levels of granularity of the same observations, or aggregate properties of the observation sequence (Lafferty et al. 2001).

3.2 Representation

3.2.1 Corpora for word embedding generation

A crucial issue to get a robust system is the means in which the data were represented. In this work, we explored alternative characterizations ranging from classical discrete representations to state-of-the-art continuous representations. The intuition is that the continuous representations based on embeddings can help to overcome the data sparsity and the lexical variability of the EHRs, improving the detection of negated entities. The embeddings are able to group semantically related words, and their main benefit is the ability to generalize over unseen words (Goldberg and Hirst 2017). Specifically, we used two types of embeddings obtained from different corpora written in Spanish: (i) in-domain corpus and (ii) general-domain corpus. The in-domain corpus comprises unannotated EHRs (denoted as uEHRs) from the same hospital to those in the IxaMed-GS. Nevertheless, this dataset does not include the EHRs used to train the CRFs classifier or to evaluate it. The general-domain corpus is the Spanish Billion Words Corpus and Embeddings (SBWCE) (Cardellino 2016), which contains medical texts from the IULA Treebank among others. The EHRs of the IULA clinical corpus used in the negation detection are not included. Details of the datasets are summarized in Table 1. Note that the number of tokens in SBWCE is approximately 1000 times higher than in uEHRs, but the vocabulary is just four times. The in-domain and general-domain corpora are different in the sense that with the in-domain corpus, medical term embeddings can be obtained, while more robust embeddings are provided with the out-domain corpus as there are more documents in the training process (as shown in Table 1).

Table 1 Unannotated datasets used to infer embedded representations

Dataset	Domain	Tokens	Vocabulary
uEHRs	In-	107,884,773	286,984
SBWCE	General-	1,416,342,219	1,000,653

3.2.2 Features

The features used to represent the tokens (excluding the punctuation marks) in the different experiments comprise: word-based representations ('Words'), the embedding-based representation ('Embeddings'), the clusters derived from the embeddings ('Cluster'), the entity labels ('Entity') and the negation trigger word labels ('Trigger'). These are the descriptions of each feature:

- *Words*
  - *Word form*: Surface form of the tokens (that might encompass multi-word forms) derived from the Free Ling-Med analyzer (Oronoz et al. 2013).
- *Embeddings*
  - *uEHRs*: Vector corresponding to each word form created with the in-domain corpus.
  - *SBWCE*: Vector corresponding to each word form created with the general-domain corpus.
- *Clusters*
  - *K-means*: Cluster assigned to each word using the k-means algorithm, which assigns the vector to the cluster with the nearest centroid.
  - *Brown*: Cluster assigned to each word using the Brown algorithm, which merges those clusters for which the loss in the average mutual information is least.
  - *Brown truncated*: Aforementioned Brown cluster truncated to reduce the granularity.
- *Entity*
  - *Manual*: Label in BIO notation indicating whether the token belongs (BI) or not (O) to a medical entity according to the annotations made by the experts.
  - *CRFs*: Label in BIO notation indicating whether the token belongs (BI) or not (O) to a medical entity according to the predictions made by a CRFs classifier.
- *Trigger*
  - *Label*: Label in BIO notation indicating whether the token belongs (BI) or not (O) to a negation trigger word. For example, for the word 'no,' we use the label 'B-Neg_TriggerWord.'

Table 2 Features used for each characterization employed to represent the documents

	Words	Embeddings		Clusters			Entity		Trigger	
Characterization	Word form	uEHRs	SBWCE	K-means	Brown	Brown truncated	Manual	CRFs	Label	Dimension
B1	✓									1
B2		✓	✓							10
C2.1		✓	✓					✓	✓	12
C2.2		✓	✓	✓	✓	✓		✓	✓	15
U2.1		✓	✓				✓		✓	12
U2.2		✓	✓	✓	✓	✓	✓		✓	15

The last column shows the resulting number of features

3.2.3 Characterization

The experimentation of this work focuses on assessing different characterizations created combining the features described before. The aim is to prove that the continuous representations could improve the detection of negated entities in EHRs. To this end, we created the characterizations described below. Firstly, we created the baselines. The first baseline (B1) makes use of word forms. The second baseline (B2) exploits the embedding corresponding to the word forms. In this way, we could observe whether simple embedding-based features outperformed word-based features. Secondly, we included features related to the entities, the negation trigger word and the clusters. The characterization C2.1 uses the word embedding together with the CRFs entity label and negation trigger word label. Specifically, the characterization C2.2 adds to the previous characterization the features corresponding to the clusters. These characterizations allowed us to compare again words and embeddings but, in this case, enhanced with features supposed useful for negation detection. The characterizations U2.1 and U2.2 are like the previous ones but using the gold entities in order to have an upper threshold. Table 2 summarizes the alternative characterizations together with the total number of features employed.

4 Results

4.1 Corpora

Our experiments were conducted on two corpora that comprise EHRs in Spanish annotated by experts: IxaMed-GS (Oronoz et al. 2015) and IULA (Marimon et al. 2017). They were divided into train, dev and test sets in order to infer, fine-tune and evaluate the system. Table 3 shows a quantitative description of both corpora. For each subset, the number of documents, words (words of the documents, including the repetitions) and vocabulary (words of the document, without repetitions), non-negated (⊕Ent) and negated (⊖Ent) entities are given.

Table 3 Description of IxaMed-GS and IULA corpora

Corpus	Subset	Doc	Word	Voc	⊕Ent	⊖Ent
IxaMed-GS	Train	41	20,689	1,934	1,444	400
	Dev	17	11,246	3,934	869	214
	Test	17	9,698	2,889	832	151
IULA	Train	3	13,937	3,168	135	606
	Dev	2	13,974	3,436	110	285
	Test	2	10,297	3,236	108	209

The number of documents (‘Doc’), word forms (‘Word’) and vocabulary (‘Voc’) in each set is provided together with the number of non-negated (⊕Ent) and negated (⊖Ent) entities

In IxaMed-GS, there are four types of entities: ‘Grp_Enfermedad’ (diseases), ‘Alergia’ (allergies), ‘Grp_Medicamento’ (drugs) and ‘Procedimiento’ (procedures). In IULA, there are four different types of entities: BODY (body parts), DISO (disorders), PROC (procedures) and SUBS (substances). It is important to explain that in the IxaMed-GS corpus, the 10% of the entities are discontinuous or formed by strings situated in different intervals of the text; this supposes a 30% of the negated entities. The presence of discontinuous entities makes the detection of negated entities more challenging.

4.2 Experimental setup

Let us give experimental details about the setup of our experiments. The embeddings of the in-domain corpus were inferred with GloVe (Pennington et al. 2014) and the embeddings of the general-domain corpus with skipNgram (Ling et al. 2015). In both cases, a window of size 10 was used and vectors of 300 components were yielded. Note that these embeddings were trained using a CPU.

The multi-word entities (formed by n words) were represented by the centroid c of the vectors involved $\mathbf{x}^{(i)}$, as is shown in (1), in order to use the average of these vectors (in a similar way that it is done in k-means clustering to group similar vectors). In this way, external resources are not needed and this arithmetic can be used for embeddings of any domain. For example, the disease w 'descompensación cardiaca izquierda' (left heart failure) comprises three words: $w^{(1)}$ = 'descompensación,' $w^{(2)}$ = 'cardiaca' and $w^{(3)}$ = 'izquierda' that correspond to three vectors: $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$ and $\mathbf{x}^{(3)}$, respectively. Then, w would be represented by $x = \frac{\mathbf{x}^{(1)}+\mathbf{x}^{(2)}+\mathbf{x}^{(3)}}{3}$.

$$\mathbf{c} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}^{(\mathbf{i})} \tag{1}$$

The words out of the embedding table were represented by the null vector $\mathbf{0}$.

In both embedding-based characterizations, the reduction of their dimensions to five components was done by means of the principal component analysis (PCA), using the Weka libraries (Hall et al. 2009). Regarding the clusters, for the k-means cluster, we used the algorithm implemented in word2vec (Mikolov et al. 2013), for the Brown cluster we used the Brown algorithm (Brown et al. 1992) and the truncation was done using a maximum of ten bits. To infer the negation detection models, we employed freely available implementations of CRFs, to be precise, CRF++ (Kudo 2005).

4.3 Evaluation

The evaluation was done using the holdout evaluation scheme: The model was trained on the training set and on a first stage evaluated on the development set. It was on the basis of the scores achieved on the development set that we fine-tuned the templates. To be precise, the template chosen was a window [−2, −1, 0,1,2] for the EHRs embeddings, [−3, −2, −1,0,1,2,3] for the SBWCE embeddings, [−2, −1,0,1,2] for the entity label, [-1,1] for the negation trigger word and [−1, 0, 1] for the clusters. With this template, a second model was trained making use of both train and dev sets, and next it was evaluated on the test set. The results obtained with the CRFs classifier in each experiment are shown in Table 4. Note that each experiment corresponds to one of the characterizations given in Table 2.

The performance of the system was evaluated using the regular metrics: precision, recall and f-measure. To compare these results, we focused on the f-measure given that we did not apply multiple-criteria decision making (MCDM) approaches (Kou et al. 2014, 2012). The system was assessed at two levels using the software of the SemEval task 'Analysis of Clinical Text' (Nakov and Zesch 2014):

Table 4 Precision (P), Recall (R) and F-measure (F) for the IxaMed-GS and IULA corpora

	Exact			Partial		
	P	*R*	*F*	*P*	*R*	*F*
(a) Results for the IxaMed-GS corpus						
B1	37.2	23.2	28.6	88.3	59.7	71.2
B2	42.0	24.5	31.0	92.0	58.3	71.4
C2.1	43.2	27.2	33.3	91.6	61.7	73.7
C2.2	42.9	27.8	33.7	91.8	63.8	75.3
U2.1	51.7	41.1	45.8	93.3	82.4	87.5
U2.2	50.8	40.4	45.0	93.3	83.0	87.8
(b) Results for the IULA corpus						
B1	78.7	58.3	67.0	92.6	68.5	78.7
B2	76.2	63.0	69.0	90.5	74.8	81.9
C2.1	77.5	62.2	69.0	90.2	72.4	80.3
C2.2	79.2	63.0	70.2	92.1	73.2	81.6
U2.1	74.6	83.5	78.8	82.4	94.4	88.0
U2.2	79.1	83.5	81.2	86.6	93.5	89.9

The evaluation was done using the train and dev sets for training and the test set for evaluation

– *Exact match*: The entity found by the system is the same as the entity annotated by the experts.
– *Partial match*: The entity found by the system and the entity of the manual annotation overlap.

4.4 Discussion

First, we compared the results obtained with word-based features (baseline B1) and embedding-based features (baseline B2). These results show that just replacing the discrete characterization by a continuous characterization, let us to improve the performance. The improvement of the f-measure for exact match and partial match is approximately a 2% for the IxaMed-GS and the IULA corpora.

After observing that the word embeddings could be useful in the characterization for the detection of negated entities, we explored the incorporation of the entity and negation trigger labels in the embedding-based characterization (C2.1). On the one hand, we can see that with IxaMed-GS, this experiment outperforms both baselines. On the other hand, with IULA we do not see an improvement in relation to the embedding-based baseline (B2), but it still outperforms the word-based baseline (B1).

Furthermore, the incorporation of the clusters created from the embeddings (characterization C2.2) can help to improve the negation detection of the baseline (baseline B2) by approximately 2% in the majority of the cases. It could be because of the incorporation of different granularity levels in the representation.

Regarding the upper threshold, we can see that with the manual annotations the f-measure is always higher than with the CRFs predictions, as it was expected. It happens because the *f*-measure for the entity recognition is 64.2 for the exact match and 86.4 for the partial match in the case of IxaMed-GS, and 60.5 for the exact match and 73.5 for the partial match in the case of IULA. The results show that for the IxaMed-GS corpus, the f-measure obtained with the entities predicted by the CRFs classifier gets worse approximately a 12% the f-measure of the upper threshold. For the IULA corpus, the f-measure gets worse approximately a 9%.

In addition, we saw that if we do not take into account the discontinuous entities during the evaluation, the f-measure increases approximately a 7% in the exact match and gets worse a 7% in the partial match (10% in the upper threshold). This reflects that the discontinuous entities are only found partially. We also observed that this system can be able to detect entities that appear negated by means of a prefix, for instance, 'asintomático' (asymptomatic).

5 Conclusions and future work

The aim of this work is to create a system for the detection of negated entities in medical texts. The main characteristic is the use of a continuous characterization to represent the documents instead of the symbolic word-based features. We proved that as in other tasks such as NER, the use of word embeddings is helpful for detecting negated entities. Above all, the word embeddings are useful in tasks with sparse data, as happens with the medical documents. To the best of our knowledge, this is the first time that negation detection was done using embeddings in the characterization for texts written in Spanish.

We envisaged negation detection for a real application to extract adverse drug reactions (ADRs), that is, the extraction of causal relations between drug and disease entities. Negated entity recognition is a key issue for early discarding absence of ADRs and thus to prevent the relation extraction system from analyzing them.

As future work, we could test whether other supervised classifiers such as SVM would obtain more benefit from the embedding-based characterization. That is to say, we would use the same characterizations with the SVM classifier. Apart from that, we could include clusters obtained with the SBWCE corpus to the representation. The motivation would be to see whether clusters generated with general-domain texts can help to improve the results also in the upper-threshold experiments.

Acknowledgements This work was partially funded by the Spanish Ministry of Science and Innovation (PROSAMED: TIN2016-77820-C3-1-R and TADEEP: TIN2015-70214-P) and the Basque Government (DETEAMI: Ministry of Health 2014111003, Predoctoral Grant: PRE 2016 1 0128).

Compliance with ethical standard

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Agarwal S, Yu H (2010) Biomedical negation scope detection with conditional random fields. J Am Med Inform Assoc 17(6):696–701

Artetxe M, Labaka G, Lopez-Gazpio I, Agirre E (2018) Uncovering divergent linguistic information in word embeddings with lessons for intrinsic and extrinsic evaluation. In: Proceedings of the 22nd conference on computational natural language learning, pp 282–291

Asker L, Boström H, Papapetrou P, Persson H (2016) Identifying factors for the effectiveness of treatment of heart failure: a registry study. In: 2016 IEEE 29th international symposium on computer-based medical systems, pp 205–206

Brown PF, deDouza PV, Mercer RL, Della Pietra VJ, Lai JC (1992) Class-based n-gram models of natural language. Computat Linguist 18(4):467–479

Cardellino C (2016) Spanish billion words corpus and embeddings. http://crscardellino.me/SBWCE/

Casillas A, Pérez A, Oronoz M, Gojenola K, Santiso S (2016) Learning to extract adverse drug reaction events from electronic health records in Spanish. Expert Syst Appl 61:235–245

Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG (2001) A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 34(5):301–310

Cohen G, Hilario M, Sax H, Hugonnet S, Geissbuhler A (2006) Learning from imbalanced data in surveillance of nosocomial infection. Artif Intell Med 37(1):7–18

Copara J, Ochoa J, Thorne C, Glavaš G (2016) Spanish NER with word representations and conditional random fields. In: Proceedings of the sixth named entity workshop, pp 34–40

Costumero R, López F, Gonzalo-Martín C, Millan M, Menasalvas E (2014) An approach to detect negation on medical documents in Spanish. Int Conf Brain Inform Health 8609:366–375

Cotik V, Roller R, Xu F, Uszkoreit H, Budde K, Schmidt D (2016) Negation detection in clinical reports written in German. In: Proceedings of the fifth workshop on building and evaluating resources for biomedical text mining, pp 115–124

Cruz NP, Maña MJ, Mata J (2010) Aprendizaje automático versus expresiones regulares en la detección de la negación y la especulación en biomedicina. Proces Leng Nat 45:77–85

Cruz NP, Maña MJ, Mata J (2012) A machine-learning approach to negation and speculation detection in clinical texts. J Assoc Inf Sci Technol 63(7):1398–1410

Cruz NP, Morante R, Maña MJ, Mata J, Parra CL (2017) Annotating negation in Spanish clinical texts. In: Proceedings of the workshop computational semantics beyond events and roles, pp 53–58

Dalianis H, Névéol A, Savova G, Zweigenbaum P (2014) Didactic panel: clinical natural language processing in languages other than English. AMIA Annu Symp 2014:1–12

Deléger L, Grouin C (2012) Detecting negation of medical problems in French clinical notes. In: Proceedings of the 2nd ACM sighit international health informatics symposium, pp 697–702
Goldberg Y, Hirst G (2017) Neural network methods in natural language processing. Morgan & Claypool Publishers, San Rafael
Goldin IM, Chapman WW (2003) Learning to detect negation with 'not' in medical texts. In: SIGIR 2003 workshop on text analysis and search for bioinformatics, pp 1–7
Gormley MR, Yu M, Dredze M (2015) Improved relation extraction with feature-rich compositional embedding models. In: Proceedings of the 2015 conference on empirical methods in natural language processing, pp 1774–1784
Guo J, Che W, Wang H, Liu T (2014) Revisiting embedding features for simple semi-supervised learning. In: Proceedings of the 2014 conference on empirical methods in natural language processing, pp 110–120
Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. SIGKDD Explor Newsl 11(1):10–18
He H, Lin J (2016) Pairwise word interaction modeling with deep neural networks for semantic similarity measurement. In: Proceedings of human language technologies: the annual conference of the North American chapter of the association for computational linguistics, pp 937–948
Henriksson A, Zhao J, Boström H, Dalianis H (2015) Modeling electronic health records in ensembles of semantic spaces for adverse drug event detection. In: 2015 IEEE international conference on bioinformatics and biomedicine, pp 343–350
Jacobson O, Dalianis H (2016) Applying deep learning on electronic health records in Swedish to predict healthcare-associated infections. In: Proceedings of the 15th workshop on biomedical natural language processing, pp 191–195
Kang T, Zhang S, Xu N, Wen D, Zhang X, Lei J (2017) Detecting negation and scope in Chinese clinical notes using character and word embedding. Comput Methods Programs Biomed 140:53–59
Kou G, Lu Y, Peng Y, Shi Y (2012) Evaluation of classification algorithms using mcdm and rank correlation. Int J Inf Technol Decis Mak 11(01):197–225
Kou G, Peng Y, Wang G (2014) Evaluation of clustering algorithms for financial risk analysis using mcdm methods. Inf Sci 275:1–12
Kudo T (2005) CRF++: Yet another CRF toolkit. https://sourceforge.net/projects/crfpp/
Lafferty J, McCallum A, Pereira F (2001) Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of the eighteenth international conference on machine learning, vol 1, pp 282–289
Ling W, Dyer C, Black AW, Trancoso I (2015) Two/too simple adaptations of word2vec for syntax problems. In: Proceedings of the 2015 conference of the North American chapter of the association for computational linguistics: human language technologies, pp 1299–1304
Marimon M, Vivaldi J, Bel N (2017) Annotation of negation in the IULA Spanish clinical record corpus. In: Proceedings of the workshop computational semantics beyond events and roles, pp 43–52
Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. In: Proceedings of workshop at international conference on learning representations, pp 1–12
Morante R, Daelemans W (2009) A metalearning approach to processing the scope of negation. In: Proceedings of the thirteenth conference on computational natural language learning, pp 21–29
Nakov P, Zesch T (eds) (2014) Proceedings of the 8th international workshop on semantic evaluation
Nguyen TH, Grishman R (2016) Combining neural networks and log-linear models to improve relation extraction. In: Proceedings of IJCAI workshop on deep learning for artificial intelligence, pp 1–7
Oronoz M, Casillas A, Gojenola K, Pérez A (2013) Automatic annotation of medical records in Spanish with disease, drug and substance names. In: Progress in pattern recognition, image analysis, computer vision, and applications—18th Iberoamerican congress, vol 8259, pp 536–543
Oronoz M, Gojenola K, Pérez A, Díaz de Ilarraza A, Casillas A (2015) On the creation of a clinical gold standard corpus in Spanish: mining adverse drug reactions. J Biomed Inform 56:318–332
Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing, pp 1532–1543
Pérez A, Gojenola K, Casillas A, Oronoz M, Díaz de Ilarraza A (2015) Computer aided classification of diagnostic terms in Spanish. Expert Syst Appl 42(6):2949–2958
Segura-Bedmar I, Suárez-Paniagua V, Martínez P (2015) Exploring word embedding for drug name recognition. In: Proceedings of the sixth international workshop on health text mining and information analysis, pp 64–72
Skeppstedt M (2011) Negation detection in Swedish clinical text an adaption of NegEx to Swedish. J Biomed Semant 2(3):1–12
Vincze V, Szarvas G, Farkas R, Móra G, Csirik J (2008) The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. BMC Bioinform 9(11):1–9

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.