Supervised Learning Approaches to Detect Negation Cues in Spanish Reviews

Lluís Domínguez-Mas, Francesco Ronzano, and Laura Furlong

Integrative Biomedical Informatics group
Research Programme on Biomedical Informatics (GRIB)
Hospital del Mar Medical Research Institute (IMIM) and Universidad Pompeu Fabra
Barcelona, 08003, Spain
lluis.dominguez01@estudiant.upf.edu,{francesco.ronzano,laura.furlong}@upf.edu

Abstract. The availability of automated approaches to effectively detect and characterize negation in textual contents is essential to robustly perform a wide range of Natural Language Processing tasks. The Workshop on Negation in Spanish 2019 provides a forum to share investigations and compare methodologies dealing with the characterization of negation in Spanish texts. In this paper we present our participation to Sub-task A organized in the context of this Workshop, focusing on the detection of negation cues in Spanish product reviews. We consider four negation cues detection approaches based on supervised learning and compare their performance. The best performing approach, based on a Conditional Random Fields sequence labeller, has been evaluated in the context of the Sub-task A of the Workshop on Negation in Spanish 2019 obtaining robust performance and scoring as the most precise system across several text domains, while keeping acceptable recall rates.

Keywords: Negation Detection · Natural Language Processing · Conditional Random Fields

1 Introduction

Negation represents a core linguistic phenomenon that aims at reversing the truth value of a statement [9]. Both the detection of negations and the identification of their scope (i.e. the text excerpts where the information that is actually negated is described) constitute essential steps towards a consistent interpretation of the meaning of natural language texts across a wide range of domains. As a consequence, automated approaches to characterize negations often represent key components to support a diverse set of Natural Language Processing tasks including *Sentiment Analysis* [19, 5], *Clinical Text Mining* [24, 16], *Relation extraction* [3, 21] and *Machine Translation* [1, 23, 7]. Even if during the last few years several efforts have been done to address a wider range of languages [22, 4, 17], nowadays English still focuses most of the investigations dealing with negation detection approaches. Especially when we consider exclusively the identification of negation cues, string matching and rule-based methodologies obtain an acceptable performance across a wide range of domains. In this regard, NegEx [2] is on of the best known examples of rule-based algorithms to characterize negations, originally tailored to clinical text. More recent approaches based on NegEx have also considered the results of dependency and constituency parsing of textual contents to improve the detection of negations: among them there are DEEPEN [16] and Negation Resolution [8]. During the last few years, thanks also to the increasing availability of corpora where the occurrences of negation have been manually annotated, several methodologies to detect negation based on supervised learning have been proposed [5, 20, 14].

As in previous editions [12, 11], the Workshop on Negation in Spanish 2019 (NEGES 2019) [10] provides a forum to share investigations and compare approaches dealing with the characterization of negation in Spanish texts. In the previous edition of the NEGES Workshop (2018), two approaches have been proposed to automatically identify negation cues in the textual contents of Spanish product reviews, both based on supervised learning techniques, namely Conditional Random Fields sequence labelling [15] and Bidirectional-LSTM [6].

The participants of NEGES 2019 have been proposed two sub-tasks concerning the *detection of negation cues* (Sub-task A) and the *assessment of the role of negation in sentiment analysis* (Sub-task B). In this paper, we describe our participation to the Sub-task A of NEGES 2019, presenting our approach to detect negation cues. In particular, after providing a brief overview of the Sub-task A in Section 2, Section 3 introduces the four supervised learning approaches to negation cue detection that we have considered in our experiments: these approaches are evaluated and thus compared by relying on the train dataset of the Sub-task A of NEGES 2019. We have chosen the best performing negation detection approach to support our participation to NEGES 2019, whose official evaluation results are discussed in Section 4. To conclude, in Section 5 we present our final remarks and plans for future work.

2 The Sub-task A of NEGES 2019: negation cues detection

The Sub-task A of NEGES 2019 challenged participants to develop effective approaches to automatically identify negation cues in Spanish texts. In particular, the SFU Review SP-NEG corpus [13] has been used by NEGES 2019 organizers to provide participants with manually annotated examples of negation cues, thus supporting the creation of both the train and test datasets of the Sub-task A. At time of writing the gold standard annotations of negation cues in the test dataset have not been released by the organizers of NEGES 2019.

The SFU Review SP-NEG corpus includes the text of 400 Spanish review gathered from the web portal Ciao.es and dealing with the following eight domains: movies, books, cell phones, music, hotels, cars, washing machines and computers. Table 1 provides an overview of the number of sentences by domain included in NEGES 2019 train and test datasets as well as the number of annotated negation cues that are present in the train dataset. From Table 1, we can notice that about 81% of negation cues of the train dataset span over a single token. Moreover, the majority of the negation cues of the train dataset (2,616 over 3,098) are expressions that span over one or more contiguous tokens. In particular, as we can see from Table 2 'no' is by far the most common negation cue with a total of 1,824 occurrences as single-token cue in the train dataset.

Table 1. Number of sentences and negation cues (in parenthesis the number of negation cues spanning over a single token is reported) in the train dataset by domain and number of sentences in the test dataset by domain. The development dataset is considered as part of the train dataset.

	TRAIN		TEST
Domain	Num. sent.	Num. neg. cues (one-token cues)	Num. sent.
movies	1,960	770 (626)	512
books	1,189	551 (447)	651
cell phones	921	444 (373)	100
music	738	286 (241)	215
hotels	708	301 (230)	145
cars	661	256 (203)	95
washing machines	650	268 (210)	250
computers	416	222 (181)	235
TOTAL:	7,243	3,098 (2,511)	2,203

Table 2. The 7 most frequent negation cues occurring as continuous and discontinuous text spans (frequency in the train dataset in parenthesis).

Continuous span	Discontinuous span
no (1,824)	no ... nada (98)
sin (224)	no ... ni (32)
ni (112)	no ... mucho (29)
nada (104)	no ... ningn/a (33)
nunca (60)	no ... muy (27)
nadie (46)	no ... para_nada (14)
tampoco (44)	no ... ni ... ni (13)

3 Negation cues detection approaches

We modelled the detection of negation cues as a token labeling task. In particular, we assigned to each token of the sentences of the NEGES 2019 train dataset one of the following three labels: B, I or O. By default, all the tokens of a sentences are labelled as O-tokens (tokens outside a negation cue), except the tokens belonging to a negation cue. If the negation cue spans over a single token, it is assigned the label B since represents the beginning of the cue. Otherwise, if a negation cue spans over two or more consecutive tokens, the first token is labelled as a B-token while the consecutive ones as I-tokens (tokens of a negation cue subsequent to the first one). In case a negation cue is composed by discontinuous text spans, each one of these spans is treated as a separate negation cue. The following sentence provides an example of the token labeling approach just described:

Ah_O no_B me_O esperaba_O nadie_B demostrando_O falta_B de_I seriedad_O ._O

We considered four supervised learning approaches and evaluate their ability to learn to predict the label (B, I or O) to assign to each token of a sentence, thus detecting the occurrence of negation cues. Independently from the supervised learning approach adopted, we represented each token by means of the following set of features:

- **Shallow textual features**:
  - number of characters of the token;
  - position of the token in the sentence, obtained by dividing the index of the token in the sentence by the total number of tokens that are present in the sentence, thus generating a number in the interval $[0, 1]$;
  - lower-cased token;
  - percentage of lowercase characters;
  - percentage of non alphabetic characters.
- **Lemma features**:
  - lemma;
  - if the lemma includes more than one character, first two characters and last two characters of the lemma. For instance in case of the lemma *create*, two additional textual features are created: *cr* and *te*.
- **Part of Speech features**:
  - Part of Speech category of the token (one nominal value among adjective, conjunction, determiner, etc.);
  - the complete result of the morphological analysis of the token including information about person, gender, number when appropriate. For instance the label NCFP for a noun (N), proper (P), feminine (F), plural (P).
- **Dependency tree features**: For the considered dependency tree node (i.e. token) and, if any, its parent node we considered the following features:
  - token;
  - lemma;
  - Part of Speech category (one nominal value among adjective, conjunction, determiner, etc.);
  - depth of the token in the dependency tree;

- number of children and descendent nodes;
- dependency relation towards the parent, if any.

We relied on the open-source language analysis framework Freeling (version 4.1) [18] in order to carry out the linguistic analyses needed to extract the token features just described by performing morphological analysis and dependency parsing.

By exploiting the set of features just described to characterize each token, we evaluated the performance of the following four token labelling / classification approaches:

– **Conditional Random Fields** (CRF), a sequence labelling statistical modelling method. We relied on the crfsuite[1] implementation of CRF.
– **Random Forest** (RF), an ensamble learning method for classification based on decision trees. We relied on the Random Forest implementation provided by SciKit learn[2].
– **Support Vector Machine with linear kernel** (SVM-linear), handling multi-class classification by means of the one-vs-the-rest scheme. We relied on the linear SVM implementation provided by SciKit learn.
– **XGBoost** (XGB), optimized gradient boosted decision trees. We relied on the XGBoost python package[3].

We evaluated the previous token labelling / classification approaches by considering the default values of their parameters as specified by each algorithm implementation considered. When we applied the RF, SVM-linear and XGB approaches, we characterized each token by relying on the set of features previously mentioned in order to describe both the token and all the tokens occurring in a $[-3, 3]$ window centered on that token. In this way, also information modelling the context of a token can be considered to predict the most likely label to assign to it.

Table 3 shows the performance of the four token labelling / classification approaches considered with respect to a 10-fold cross-validation over the NEGES 2019 train dataset. This Table evaluates also a BASELINE negation cue detection strategy (last column): the negation detection approach of this strategy creates a list of lemmatized negation cues from the train dataset of each fold and marks the occurrences of these cues in the test dataset. From Table 3 we can notice that the CRF is the best performing approach, with a more sensible improvement in performance when we consider the macro F-score of the BIO labels and strict matches of predicted negation cues with a gold standard ones. We have to notice that the BASELINE negation detection strategy, based on simple string match, obtains acceptable performance. Anyway, looking into further details, even if this trend is not evident when we consider the F-scores, the results of the BASELINE negation detection strategy are the ones that present

[1] http://www.chokkan.org/software/crfsuite/
[2] https://scikit-learn.org/
[3] https://xgboost.readthedocs.io/

the strongest differences among precision and recall: low values of precision are balanced by high values of recall. The other approaches based on supervised learning obtain a better balance among precision and recall.

By inspecting the types of classification errors of the four supervised learning approaches we considered, we can spot the following trend: the performance of most approaches drastically decreases when they deal with the identification of multi-token negation cues and, in particular, when the tokens that occur after the first one should be spotted. This trend could be related to the greater difficulty in characterizing linguistically these tokens and the low number of annotated examples of multi-token negation cues that are available in the NEGES 2019 corpus. In future investigations we plan to analyze in detail this issue by eventually proposing a negation cue detection strategy tailored to improve the performance of the detection of multi-token cues.

Table 3. F-score of 10-fold cross-validation over the NEGES 2019 train dataset. Evaluation approaches: *BIO*, macro F-score by considering the three labels BIO assigned to each token; *Cue marker strict*: F-score of negation cue annotations by considering as true positives only the exact matches of a predicted negation cue and a gold standard one; *Cue marker non-strict*: F-score of negation cue annotations by considering as true positives both the exact and partial matches of a predicted negation cue and a gold standard one.

Evaluation approach	CRF	RF	SVM-linear	XGB	BASELINE
BIO	0.7337	0.6220	0.5173	0.7273	0.7060
Cue marker strict	0.9150	0.9035	0.8877	0.9082	0.8379
Cue marker non-strict	0.9259	0.9228	0.9069	0.9252	0.8556

4 NEGES 2019 evaluation

We chose the best performing supervised learning approach resulting from the experiments described in Section 3 (i.e. the CRF sequence labeller) as the methodology exploited to generate our negation cue predictions for the Sub-task A of NEGES 2019. Our approach scored second in terms of precision and third in terms of recall and F-score. In particular, when we look at the negation detection results across each single domain of the text of the test set of the Sub-task A of NEGES 2019, our approach obtained the highest precision in four domains over eight (movies, mobiles, washing machines and hotels).

5 Conclusion

In this paper we presented the negation cues detection approach we devised in the context of our participation to the Sub-task A of the Workshop on Negation in Spanish 2019, dealing with the automated identification of negation cues in Spanish texts. We described in detail the four supervised learning approaches we considered for our participation to NEGES 2019, by comparing their performance on the train dataset of the Sub-task A. We also discussed the results of negation cues detection approach that we selected to participate to NEGES 2019, based on a Conditional Random Field sequence labeller. As future work, we would like to evaluate the performance of a wider range of negation cues detection systems, considering both sequence labeller based on neural network architectures, relying on word embeddings and ensebling methods that combine the predictions of distinct sequence labelling approaches. We plan also to perform a more detailed error analysis to better characterize and try to mitigate the weknesses of the negation cues detection approaches considered.

References

1. Baker, K., Bloodgood, M., Dorr, B.J., Callison-Burch, C., Filardo, N.W., Piatko, C., Levin, L., Miller, S.: Modality and negation in SIMT: use of modality and negation in semantically-informed syntactic MT. Computational Linguistics **38**(2), 411–438 (2012)
2. Chapman, W.W., Bridewell, W., Hanbury, P., Cooper, G.F., Buchanan, B.G.: A simple algorithm for identifying negated findings and diseases in discharge summaries. Journal of biomedical informatics **34**(5), 301–310 (2001)
3. Chowdhury, M.F.M., Lavelli, A.: Exploiting the scope of negations and heterogeneous features for relation extraction: a case study for drug-drug interaction extraction. In: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. pp. 765–771 (2013)
4. Cotik, V., Roller, R., Xu, F., Uszkoreit, H., Budde, K., Schmidt, D.: Negation detection in clinical reports written in German. In: Proceedings of the Fifth Workshop on Building and Evaluating Resources for Biomedical Text Mining (BioTxtM2016). pp. 115–124 (2016)
5. Cruz, N.P., Taboada, M., Mitkov, R.: A machine-learning approach to negation and speculation detection for sentiment analysis. Journal of the Association for Information Science and Technology **67**(9), 2118–2136 (2016)
6. Fabregat, H., Araujo Serna, L., Martínez Romo, J.: Deep learning approach for negation trigger and scope recognition pp. 43–48 (2018)
7. Fancellu, F., Webber, B.: Translating negation: A manual error analysis. In: Proceedings of the Second Workshop on Extra-Propositional Aspects of Meaning in Computational Semantics (ExProM 2015). pp. 2–11 (2015)
8. Gkotsis, G., Velupillai, S., Oellrich, A., Dean, H., Liakata, M., Dutta, R.: Don't Let Notes Be Misunderstood: A Negation Detection Method for Assessing Risk of Suicide in Mental Health Records. In: Proceedings of the Third Workshop on Computational Lingusitics and Clinical Psychology. pp. 95–105 (2016)
9. Horn, L.: A natural history of negation (1989)
10. Jiménez-Zafra, S.M., Cruz Díaz, N.P., Morante, R., Martín-Valdivia, M.T.: NEGES 2019 Task: Negation in Spanish. In: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2019). CEUR Workshop Proceedings, CEUR-WS, Bilbao, Spain (2019)

11. Jiménez-Zafra, S.M., Dıaz, N.P.C., Morante, R., Martın-Valdivia, M.T.: Tarea 2 del Taller NEGES 2018: Detección de Claves de Negación. In: Proceedings of NEGES 2018: Workshop on Negation in Spanish. vol. 2174, pp. 35–41 (2018)
12. Jiménez-Zafra, S.M., Díaz, N.P.C., Morante, R., Martín-Valdivia, M.T.: NEGES 2018: Workshop on Negation in Spanish. Procesamiento del Lenguaje Natural **62**, 21–28 (2019)
13. Jiménez-Zafra, S.M., Taulé, M., Martín-Valdivia, M.T., Ureña-López, L.A., Martí, M.A.: SFU Review SP-NEG: a Spanish corpus annotated with negation for sentiment analysis. A typology of negation patterns. Language Resources and Evaluation **52**(2), 533–569 (2018)
14. Lazib, L., Zhao, Y., Qin, B., Liu, T.: Negation scope detection with recurrent neural networks models in review texts. In: International Conference of Pioneering Computer Scientists, Engineers and Educators. pp. 494–508. Springer (2016)
15. Loharja, H., Padró, L., Turmo Borras, J.: Negation cues detection using CRF on Spanish product review texts. In: NEGES 2018: Workshop on Negation in Spanish: Seville, Spain: September 19-21, 2018: proceedings book. pp. 49–54 (2018)
16. Mehrabi, S., Krishnan, A., Sohn, S., Roch, A.M., Schmidt, H., Kesterson, J., Beesley, C., Dexter, P., Schmidt, C.M., Liu, H., et al.: DEEPEN: A negation detection system for clinical text incorporating dependency relation into NegEx. Journal of biomedical informatics **54**, 213–219 (2015)
17. Névéol, A., Dalianis, H., Velupillai, S., Savova, G., Zweigenbaum, P.: Clinical natural language processing in languages other than English: opportunities and challenges. Journal of biomedical semantics **9**(1), 12 (2018)
18. Padró, L., Stanilovsky, E.: Freeling 3.0: Towards wider multilinguality. In: LREC2012 (2012)
19. Pröllochs, N., Feuerriegel, S., Neumann, D.: Enhancing sentiment analysis of financial news by detecting negation scopes. In: 2015 48th Hawaii International Conference on System Sciences. pp. 959–968. IEEE (2015)
20. Qian, Z., Li, P., Zhu, Q., Zhou, G., Luo, Z., Luo, W.: Speculation and negation scope detection via convolutional neural networks. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. pp. 815–825 (2016)
21. Sanchez-Graillet, O., Poesio, M.: Negation of protein–protein interactions: analysis and extraction. Bioinformatics **23**(13), i424–i432 (2007)
22. Skeppstedt, M.: Negation detection in Swedish clinical text: An adaption of NegEx to Swedish. In: Journal of Biomedical Semantics. vol. 2 (3), p. S3. BioMed Central (2011)
23. Wetzel, D., Bond, F.: Enriching parallel corpora for statistical machine translation with semantic negation rephrasing. In: Proceedings of the Sixth Workshop on Syntax, Semantics and Structure in Statistical Translation. pp. 20–29. Association for Computational Linguistics (2012)
24. Wu, S., Miller, T., Masanz, J., Coarr, M., Halgrim, S., Carrell, D., Clark, C.: Negations not solved: generalizability versus optimizability in clinical natural language processing. PloS one **9**(11), e112774 (2014)