TASS 2016: Workshop on Sentiment Analysis at SEPLN, septiembre 2016, pág. 13-21

Overview of TASS 2016

Resumen de TASS 2016

Miguel Ángel García Cumbreras[1], Julio Villena Román[2], Eugenio Martínez Cámara[1], Manuel Carlos Díaz Galiano[1], M. Teresa Martín Valdivia[1], L. Alfonso Ureña López[1]
[1]Universidad de Jaén
23071 Jaén, Spain
[2]Sngular
28034 Madrid, Spain
[1]{magc, emcamara, mcdiaz, laurena, maite}@ujaen.es [2]{julio.villena}@sngular.team

Resumen: Este artículo describe la quinta edición del taller de evaluación experimental TASS 2016, enmarcada dentro del Congreso Internacional SEPLN 2016. El principal objetivo de TASS es promover la investigación y el desarrollo de nuevos algoritmos, recursos y técnicas para el análisis de sentimientos en medios sociales (concretamente en Twitter), aplicado al idioma español. Este artículo describe las tareas propuestas en TASS 2016, así como el contenido de los corpus utilizados, los participantes en las distintas tareas, los resultados generales obtenidos y el análisis de estos resultados.
Palabras clave: TASS 2016, análisis de opiniones, medios sociales

Abstract: This paper describes TASS 2016, the fifth edition of the Workshop on Sentiment Analysis at SEPLN. The main aim is the promotion of the research and the development of new algorithms, resources and techniques on the field of sentiment analysis in social media (specifically Twitter) focused on the Spanish language. This paper presents the TASS 2016 proposed tasks, the description of the corpora used, the participant groups, the results and analysis of them.
Keywords: TASS 2016, sentiment analysis, social media.

1 Introduction

TASS is an experimental evaluation workshop, a satellite event of the annual SEPLN Conference, with the aim to promote the research on Sentiment Analysis in social media focused on the Spanish language. The fifth edition will be held on September 13th, 2016 at the University of Salamanca, Spain.

Sentiment Analysis (SA) is traditionally defined as the computational treatment of opinion, sentiment and subjectivity in texts (Pang & Lee, 2008). However, Cambria and Hussain (2012) offer a more updated definition: Computational techniques for the extraction, classification, understanding and evaluation of opinions and comments published on the Internet and other kind of user generated contents. It is a hard task because even humans often disagree on the polarity of a given text. And it is a harder task when the text has only 140 characters (Twitter messages or tweets).

Although SA is not a new task, it is still challenging, because the state of the art has not yet resolved some problems related to multilingualism, domain adaptation, text genre adaptation and polarity classification at fine grained level. Polarity classification has usually been tackled following two main approaches. The first one applies machine learning algorithms in order to train a polarity classifier using a labelled corpus (Pang et al. 2002). This approach is also known as the supervised approach. The second one is known as semantic orientation, or the unsupervised approach, and it integrates linguistic resources in a model in order to identify the valence of the opinions (Turney 2002).

The aim of TASS is to provide a competitive forum where the newest research works in the field of SA in social media, specifically focused on Spanish tweets, are described and discussed by scientific and business communities.

The rest of the paper is organized as follows. Section 2 describes the different corpus provided to participants. Section 3 shows the different tasks of TASS 2016. Section 4 describes the participants and the overall results are presented in Section 5. Finally, the last section shows some conclusions and future directions.

2 Corpus

TASS 2016 experiments are based on two corpora, specifically built for the different editions of the workshop.

The two corpora will be made freely available to the community after the workshop. Please send an email to tass@sngularmeaning.team filling in the TASS Corpus License agreement with your email, affiliation (institution, company or any kind of organization) and a brief description of your research objectives, and you will be given a password to download the files in the password protected area. The only requirement is to include a citation to a relevant paper and/or the TASS website.

2.1 General corpus

The General Corpus contains over 68.000 tweets, written in Spanish, about 150 well-known personalities and celebrities of the world of politics, economy, communication, mass media and culture, between November 2011 and March 2012. Although the context of extraction has a Spanish-focused bias, the diverse nationality of the authors, including people from Spain, Mexico, Colombia, Puerto Rico, USA and many other countries, makes the corpus reach a global coverage in the Spanish-speaking world.

Each tweet includes its ID (*tweetid*), the creation date (*date*) and the user ID (*user*). Due to restrictions in the Twitter API Terms of Service (https://dev.twitter.com/terms/api-terms), it is forbidden to redistribute a corpus that includes text contents or information about users. However, it is valid if those fields are removed and instead IDs (including Tweet IDs and user IDs) are provided. The actual message content can be easily obtained by making queries to the Twitter API using the *tweetid*.

The general corpus has been divided into training set (about 10%) and test set (90%). The training set was released, so the participants could train and validate their models. The test corpus was provided without any tagging and has been used to evaluate the results. Obviously, it was not allowed to use the test data from previous years to train the systems.

Each tweet was tagged with its global polarity (positive, negative or neutral sentiment) or no sentiment at all. A set of 6 labels has been defined: strong positive (P+), positive (P), neutral (NEU), negative (N), strong negative (N+) and one additional no sentiment tag (NONE).

In addition, there is also an indication of the level of agreement or disagreement of the expressed sentiment within the content, with two possible values: AGREEMENT and DISAGREEMENT. This is especially useful to make out whether a neutral sentiment comes from neutral keywords or else the text contains positive and negative sentiments at the same time.

Moreover, the polarity values related to the entities that are mentioned in the text are also included for those cases when applicable. These values are similarly tagged with 6 possible values and include the level of agreement as related to each entity.

This corpus is based on a selection of a set of topics. Thematic areas such as “política” (“politics”), “fútbol” (“soccer”), “literatura” (“literature”) or “entretenimiento” (“entertainment”). Each tweet in the training and test set has been assigned to one or several of these topics (most messages are associated to just one topic, due to the short length of the text).

The annotation has been semi-automatically done: a baseline machine learning model is first run and then all tags are checked by human experts. In the case of the polarity at entity level, due to the high volume of data to check, the human annotation has only been done for the training set.

Table 1 shows a summary of the training and test corpora provided to participants.

Attribute	Value
Tweets	68.017
Tweets (test)	60.798 (89%)
Tweets (test)	7.219 (11%)
Topics	10
Users	154
Date start (train)	2011-12-02
Date end (train)	2012-04-10
Date start (test)	2011-12-02
Date end (test)	2012-04-10

Table 1: Corpus statistics

Users were journalists (*periodistas*), politicians (*políticos)* or celebrities (*famosos*). The only language involved was Spanish (*es)*.

The list of topics that have been selected is the following:

- Politics (política)
- Entertainment (entretenimiento)
- Economy (economía)
- Music (música)
- Soccer (fútbol)
- Films (películas)
- Technology (tecnología)
- Sports (deportes)
- Literature (literatura)
- Other (otros)

The corpus is encoded in XML. Figure 1 shows the information of two tweets. The first tweet is only annotated with the polarity at tweet level because there is not any entity in the text. However, the second one is annotated with the global polarity of the message and the polarity associated to each of the entities that appear in the text (UPyD and Foro Asturias).

```
<tweet>
    <tweetid>0000000000</tweetid>
    <user>usuario0</user>
    <content>
        <![CDATA['Conozco a alguien q es adicto al drama! Ja ja ja te suena d algo!]]>
    </content>
    <date>2011-12-02T02:59:03</date>
    <lang>es</lang>
    <sentiments>
        <polarity>
            <value>P+</value>
            <type>AGREEMENT</type>
        </polarity>
    </sentiments>
    <topics>
        <topic>entretenimiento</topic>
    </topics>
</tweet>
<tweet>
    <tweetid>0000000001</tweetid>
    <user>usuario1</user>
    <content>
        <![CDATA['UPyD contará casi seguro con grupo gracias al Foro Asturias.]]>
    </content>
    <date>2011-12-02T00:21:01</date>
    <lang>es</lang>
    <sentiments>
        <polarity>
            <value>P</value>
            <type>AGREEMENT</type>
        </polarity>
        <polarity>
            <entity>UPyD</entity>
            <value>P</value>
            <type>AGREEMENT</type>
        </polarity>
        <polarity>
            <entity>Foro_Asturias</entity>
            <value>P</value>
            <type>AGREEMENT</type>
        </polarity>
    </sentiments>
  - <topics>
        <topic>política</topic>
    </topics>
</tweet>
```

Figure 1: Sample tweets (General corpus)

2.2 STOMPOL corpus

STOMPOL (corpus of Spanish Tweets for Opinion Mining at aspect level about POLitics) is a corpus of Spanish tweets prepared for the research on the challenging task of opinion mining at aspect level. The tweets were gathered from 23rd to 24th of April 2015, and are related to one of the following political aspects that appear in political campaigns:

- Economics (Economía): taxes, infrastructure, markets, labour policy...
- Health System (Sanidad): hospitals, public/private health system, drugs, doctors...
- Education (Educación): state school, private school, scholarships...
- Political party (Propio_partido): anything good (speeches, electoral programme...) or bad (corruption, criticism) related to the entity
- Other aspects (Otros_aspectos): electoral system, environmental policy...

Each aspect is related to one or several entities that correspond to one of the main political parties in Spain, which are:

- Partido_Popular (PP)
- Partido_Socialista_Obrero_Español (PSOE)
- Izquierda_Unida (IU)
- Podemos
- Ciudadanos (C's)
- Unión_Progreso_y_Democracia (UPyD)

Each tweet in the corpus has been manually annotated by two annotators, and a third one in case of disagreement, with the sentiment polarity at aspect level. Sentiment polarity has been tagged from the point of view of the person who writes the tweet, using 3 levels: P, NEU and N. Again, no difference is made between no sentiment and a neutral sentiment (neither positive nor negative). Each political aspect is linked to its correspondent political party and its polarity.

Figure 2 shows the information of two sample tweets.

```
<tweet id="591267548311769088">@ahorapodemos @Pablo_Iglesias_ @SextaNocheTV
Que alguien pregunte si habrá cambios en las <sentiment aspect="Educacion"
entity="Podemos" polarity="NEU">becas</sentiment> MEC para universitarios, por
favor.</tweet>

<tweet id="591192167944736769">#Arroyomolinos lo que le interesa al ciudadano
son Políticos cercanos que se interesen y preocupen por sus problemas <
sentiment aspect="Propio_partido" entity="Union_Progreso_y_Democracia" polarity
="P">@UPyD</sentiment> VECINOS COMO TU</tweet>
```

Figure 2: Sample tweets (STOMPOL corpus)

The number of tweets per each entity are shown in Table 2.

Entity	Train	Test
PP	205	125
PSOE	136	70
C's	119	87
Podemos	98	80
IU	111	43
UPyD	97	124
Total	766	529

Table 2: Number of tweets per entity and per corpus subset

3 *Description of tasks*

Since the first edition of TASS, a new task and a new corpus have been published. However, one of the aims of TASS is the evaluation of the progress of the research on SA. Thus, the edition of 2016 was focused on the analysis and the comparison of the systems with the submissions of previous editions.

The edition of 2016 was focused on two tasks: polarity classification at tweet level and polarity classification at entity level. The polarity classification task has been proposed with the same corpus since the first edition of TASS, but the polarity classification at aspect level has been proposed with a different corpus each edition. In the edition of 2016 the classification at aspect level uses the STOMPOL corpus, which was published the first time in the edition of 2015.

Participants are expected to submit up to 3 results of different experiments for one or both of these tasks, in the appropriate format described below.

Along with the submission of experiments, participants have been invited to submit a paper to the workshop in order to describe their experiments and discussing the results with the audience in a regular workshop session.

The two proposed tasks are described next.

3.1 Task 1: Sentiment Analysis at Global Level

This task consists on performing an automatic polarity classification to determine the global polarity of each message in the test set of the General Corpus. The training set of the corpus was provided to the participants with the aim they could train and validate their models with it. There were two different evaluations: one based on 6 different polarity labels (`P+`, `P`, `NEU`, `N`, `N+`, `NONE`) and another based on just 4 labels (`P`, `N`, `NEU`, `NONE`).

Participants are expected to submit (up to 3) experiments for the 6-labels evaluation, and they are also allowed to submit (up to 3) specific experiments for the 4-labels scenario.

Results must be submitted in a plain text file with the following format:

```
tweetid \t polarity
```

where polarity can be:

- `P+`, `P`, `NEU`, `N`, `N+` and `NONE` for the 6-labels case
- `P`, `NEU`, `N` and `NONE` for the 4-labels case.

The same test corpus of previous years was used for the evaluation in order to develop a comparison among the systems. The accuracy is one of the measures used to evaluate the systems, however due to the fact that the training corpus is not totally balanced the systems were also assessed by the macro-averaged precision, macro-averaged recall and macro-averaged F1-measure.

3.2 Task 2: Aspect-based sentiment analysis

A corpus with the entities and the aspect identified was provided to the participants, so the goal of the systems is the inference of the polarity at the aspect-level. As in 2015, STOMPOL corpus was the corpus used in this task. STOMPOL was divided in training and test set, the first one for the development and validation of the systems, and the second for evaluation.

Participants are expected to submit up to 3 experiments for each corpus, each in a plain text file with the following format:

```
tweetid \t aspect-entity \t polarity
```

Allowed polarity values are: P, N and NEU. For the evaluation, a single label combining "aspect-polarity" has been considered. As in the first task, accuracy, macro-averaged precision, macro-averaged recall and macro-averaged F1-measure have been calculated for the global result.

4 *Participants and Results*

This year 7 (7 last year) groups submitted their systems The list of active participant groups is shown in Table 3, including the tasks in which they have participated.

Six of the seven participant groups sent a report describing their experiments and results achieved. Papers were reviewed and included in the workshop proceedings. References are listed in Table 4.

Group	1	2
jacerong	X	
ELiRF-UPV	X	X
LABDA	X	
INGEOTEC	X	
GASUCR	X	
GTI		X
SINAI_w2v	X	
Total	6	1

Table 3: Participant groups

Group	Report
ELiRF	ELiRF-UPV en TASS 2016: Análisis de Sentimientos en Twitter
GTI	GTI at TASS 2016: Supervised Approach for Aspect Based Sentiment Analysis in Twitter
jacerong	JACERONG at TASS 2016: An Ensemble Classifier for Sentiment Analysis of Spanish Tweets at Global Level
LABDA	LABDA at the 2016 TASS challenge task: using word embedding for the sentiment analysis task
SINAI	Participación de SINAI en TASS 2016

Table 4: Participant reports

5 *Results*

This section will be focused on the description and the analysis of the results and the systems submitted by the participants.

5.1 Task 1: Sentiment Analysis at Global Level

Submitted runs and results for Task 1, evaluation based on 5 polarity levels with the whole General test Corpus are shown in Table 5. Accuracy, macro-averaged precision, macro-averaged recall and macro-averaged F1-measure have been used to evaluate each individual label and ranking the systems.

Run Id	M-F1
ELiRF-UPV_1	0.518
jacerong_2	0.504
jacerong_3	0.503
jacerong_1	0.499
ELiRF-UPV_2	0.496
INGEOTEC	0.464
LABDA_1	0.429
LABDA_2	0.429
LABDA_3	0.418
GASURC_3	0.254
GASURC_1	0.232
GASURC_2	0.227

Table 5: Results for Task 1, 5 levels

In order to perform a more in-depth evaluation, results are calculated considering the classification only in 3 levels (`POS`, `NEU`, `NEG`) and no sentiment (`NONE`) merging `P` and `P+` in only one category, as well as `N` and `N+` in another one. The results reached by the submitted systems are shown in Table 6.

Run Id	M-F1
jacerong_3	0.568
jacerong_2	0.567
jacerong_1	0.564
ELiRF-UPV_1	0.549
ELiRF-UPV_2	0.548
INGEOTEC	0.524
LABDA_3	0.511
LABDA_2	0.508
LABDA_1	0.508
SINAI_w2v_1	0.504
SINAI_w2v_3	0.486
SINAI_w2v_4	0.469
SINAI_w2v_2	0.440
GASURC_1	0.250
GASURC_2	0.152

Table 6: Results for Task 1, 3 levels

5.2 Task 2: Aspect-based Sentiment Analysis

Submitted runs and results for Task 2, with the STOMPOL corpus, are shown in Table 7. Accuracy, macro-averaged precision, macro-averaged recall and macro-averaged F1-measure have been used to evaluate each individual label and ranking the systems.

Run Id	M-F1
ELiRF-UPV_1	0.526
GTI	0.463

Table 7: Results for Task 2

5.3 Description of the systems

The systems submitted in the edition of 2016 represent the next step of the ones submitted in the previous edition. The systems may be cluster in two groups, those ones that rely on the classification power of the ensemble of several base classifiers, and those systems that change the use traditional Bag-of-Words model for the use of vectors of word embeddings in order to represent the meaning of each word. In the subsequent paragraphs the main features of the systems submitted are going to be depicted.

Hurtado and Pla (2016) describe the participation of the team ELiRF-UPV in the two tasks of TASS 2016. The only difference between the systems submitted for the two tasks is the fact that the one focused on the second task has a module for the identification of the context of each of the entities and aspects annotated on the tweets. The polarity classification system relies on the ensemble of 192 configurations of a SVM classifiers. For the combination of the set of classifiers they evaluate the performance of an approach based on voting and other on stacking.

The system depicted in (Cerón-Guzmán, 2016) is also based on an approach of ensemble classifiers. In this case the base classifiers used a classifier based on logistic regression and they are combined by voting.

Alvarez et al. (2016) exposed the participation of the team GTI on the task 2. The system is similar to the system of the team ELiRF-UPV in the sense that it is composed by two layers: context identification and polarity classification. Regarding the identification of the context, the authors design a heuristic method based on lexical markers. The polarity classification system is a SVM classifier that uses different type of features in order to represent the contexts of the entities and the aspects.

Montejo-Ráez and Díaz-Galiano (2016) introduce a system based on a supervised learning algorithm over vectors resulting from a weighted vector. This vector is computed using a Word2Vec algorithm. This method, which is inspired from neural-network language modelling, was executed with a collection of tweets written in Spanish and the Spanish Wikipedia in order to generate a set of word embeddings for the representation of the words of the General Corpus of TASS as dense vectors. The creation of the collection of tweets written in Spanish followed a distant supervision approach by means the assumption that tweets with happy and sad emoticons express emotions or opinions. Their experiments show massive data from Twitter can lead to a slight improvement in classification accuracy.

The system presented by the team LABDA (Quirós, Segura-Bedmar and Paloma Martínez, 2016) is similar to the one submitted by SINAI (Montejo-Ráez and Díaz-Galiano, 2016) because it also used word embeddings as schema of representation of the meaning of the words of the tweets. Quirós, Segura-Bedmar and Paloma Martínez (2016) assessed the performance of the SVM and Logistic Regression as classifiers.

Casasola Murillo and Marín Reventós (2016) submitted an unsupervised system based on the system described in Turney (2002), but with a specific adaptation to the classification of tweets written in Spanish.

5.4 Analysis

In Table 5 and Table 6 are shown the results of each system and they are ranked by the F1-score reached, so it is not hard to know what is the best system in the edition of 2016.

On the other hand, how many tweets were rightly classified by the submitted systems? Is there a set of tweets that were not rightly classified by any system? What are the most difficult tweets to classify? These questions are going to be answered in the following paragraphs?

Table 8 shows the rate of tweets that are rightly classified by a number of systems. There are about a 6% of tweets whose polarity is not inferred by any of the submitted systems. In other words, the submitted systems in the edition of 2016 are able to classify about the 94% of the test set. So, what is the main features of that 6% of tweets that any system inferred their polarity?

Number of systems	Rate of tweets
0	0.056%
1	0.065%
2	0.063%
3	0.067%
4	0.059%
5	0.061%
6	0.074%
7	0.078%
8	0.081%
9	0.112%
10	0.122%
11	0.082%
12	0.062%
13	0.011%

Table 8: Rate of tweets rightly classified (6 classes) by a number of systems

Id: 171304000392663040

Sacarle 17 puntos en la final de Copa al Barça CB en el Palau Sant Jordi es una pasada.

Beating Barça by 17 points in the Copa is amazing

Polarity: P+

Figure 3: Tweet not rightly classified by any system

Figures Figure 3,Figure 4Figure 5 are three examples of tweets that were not rightly classified by any system. The common feature of the three tweets is that they do not have any lexical marker that express emotion or opinion. Moreover, the tweet of the Figure 4 is sarcastic, which means an additional challenging for SA because requires a deep understanding of the language.

Id: 177439342497767424

hahahahahaha "@Absolutexe: ¿Le han cambiado ya el nombre a la Junta de Andalucía por la Banda de Andalucía o aún no?"

hahahahahaha "@Absolutexe: Has the Junta de Andalucía renamed Gang of Andalucía or not yet?"

Polarity: N+

Figure 4: Tweet not rightly classified by any system

Id: 177439342497767424

Rubalcaba pide a Rajoy que presente ya los Presupuestos y dice que no lo hace porque espera a las elecciones andaluzas

Rubalcaba requires Rajoy to submit the Budget and says that he didn't because he is waiting the results of the elections in Andalucia

Polarity: NONE

Figure 5: Tweet not rightly classified by any system

All the systems submitted are based on linear classifiers that do not take into account the context of each word, which means a big drawback for the understanding the meaning of a span of text.

The tweets of the Figures 3, 4 and 5 show that opinions and emotions are not only expressed by lexical markers, so the future participants should take into account the challenging task of implicit opinion analysis, irony and sarcasm detection. These new problems may be framed on the semantic level of Natural Language Processing and should be tackled by the research community in order to go a step further in the understanding of the subjective information, which is continuously published on the Internet.

6 Conclusions and Future Work

TASS was the first workshop about SA focused on the processing of texts written in Spanish. In the three first editions of TASS, the research community were mainly formed by Spanish researchers, however since the last edition, the researchers that come from South America is making bigger, so it is an evidence that the research community of Sentiment Analysis in Spanish is not only located in Spain and is formed by the Spanish speaking countries.

Anyway, the developed corpus and gold standards, and the reports from participants will for sure be helpful for knowing the state of the art in SA in Spanish.

The future work will be mainly focused on the definition of a new General Corpus because of the following reasons:

1. The language used on Twitter changes faster than the language used in traditional genres of texts, so the update of the corpus is required in order to cover a real used of the language on Twitter.
2. After several editions of the workshop, we realize that the quality of the annotation is not extremely good, so it is required to define a new corpus with a high quality annotation in order to provide a real gold standard for Spanish SA on Twitter.
3. The research community deeply know the General Corpus of TASS and it wants a new challenge.

A significant amount of new tasks is currently being defined in Natural Language Processing, so some of them, such as stance classification, will be studied to be proposal for the next edition of TASS.

Acknowledgements

This work has been partially supported by a grant from the Fondo Europeo of Desarrollo Regional (FEDER) and REDES project (TIN2015-65136-C2-1-R) from the Spanish Government.

References

Cambria, E. and Amir Hussain, A. 2012. Sentic Computing. Techniques, Tools and Applications. Springer Briefs in Cognitive Computation, volume 2. Springer Netherlands. ISBN 978-94-007-5069-2. doi:10.1007/978-94-007-5070-8.

Cerón-Guzmán, J. A. 2016. JACERONG at TASS 2016: An Ensemble Classifier for Sentiment Analysis of Spanish Tweets at Global Level. In Proceedings of TASS 2016: Workshop on Sentiment Analysis at SEPLN co-located with the 32nd SEPLN Conference (SEPLN 2016), Salamanca, September

Casasola Murillo, E. and Gabriela M. R. 2016. Evaluación de Modelos de Representación del Texto con Vectores de Dimensión Reducida para Análisis de Sentimiento. In Proceedings of TASS 2016: Workshop on Sentiment Analysis at SEPLN co-located with the 32nd SEPLN Conference (SEPLN 2016), Salamanca, September

Hurtado, Ll. and Ferran P. 2016. ELiRF-UPV en TASS 2016: Análisis de Sentimientos en Twitter. In Proceedings of TASS 2016: Workshop on Sentiment Analysis at SEPLN co-located with the 32nd SEPLN Conference (SEPLN 2016), Salamanca, September

Montejo-Ráez, A. and Díaz-Galiano, M. C. 2016. Participación de SINAI en TASS 2016. In Proceedings of TASS 2016: Workshop on Sentiment Analysis at SEPLN co-located with the 32nd SEPLN Conference (SEPLN 2016), Salamanca, September

Pang, B., Lillian Lee and Shivakumar Vaithyanathan. 2002. Thumbs up?: Sentiment classification using machine learning techniques. In Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, EMNLP '02, páginas 79–86. Association for Computational Linguistics, Stroudsburg, PA, USA. doi:10.3115/1118693.1118704.

Pang, B. and Lillian Lee (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1-2):1–135. ISSN 1554-0669. doi:10.1561/1500000011.

Quirós, A., Isabel S. B. and Paloma M. 2016. LABDA at the 2016 TASS challenge task: using word embeddings for the sentiment analysis task. In Proceedings of TASS 2016: Workshop on Sentiment Analysis at SEPLN co-located with the 32nd SEPLN Conference (SEPLN 2016), Salamanca, September

Turney, P. D. 2002. Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, ACL '02, pp: 417–424. Association for Computational Linguistics, Stroudsburg, PA, USA. doi:10.3115/1073083.1073153.

Villena-Román, J., Sara, L. S., Eugenio M. C., and José Carlos G. C. 2013. *TASS - Workshop on Sentiment Analysis at SEPLN*. Revista de Procesamiento del Lenguaje Natural, 50, pp 37-44.

Villena-Román, J., Janine G. M., Sara L. S. and José Carlos G. C. 2014. *TASS 2013 - A Second Step in Reputation Analysis in Spanish*. Revista de Procesamiento del Lenguaje Natural, 52, pp 37-44.