TASS 2017: Workshop on Semantic Analysis at SEPLN, septiembre 2017, págs. 13-21

Overview of TASS 2017

Resumen de TASS 2017

Eugenio Martínez-Cámara[1], **Manuel C. Díaz-Galiano**[2], **M. Ángel García-Cumbreras**[2], **Manuel García-Vega**[2], **Julio Villena-Román**[3]
[1]Ubiquitous Knowledge Processing Lab (UKP-TUDA)
Department of Computer Science, Technische Universität Darmstadt
[2]Grupo de Investigación SINAI
Universidad de Jaén, Jaén, Spain
[3]MeaningCloud, Madrid, Spain
[1]camara@ukp.informatik.tu-darmstadt.de,
[2]{mcdiaz, magc, mgarcia}@ujaen.es, [3]julio.villena@sngular.team

Abstract: This paper describes TASS 2017, the sixth edition of the Workshop on Semantic Analysis at SEPLN 2017. The main aim is to encourage the research and development of new resources, algorithms and techniques for different tasks of semantic analysis in Spanish. In this paper, we present the proposed tasks, the generated datasets, and a summary of the submitted systems.
Keywords: TASS 2017, sentiment analysis, semantic analysis

Resumen: Este artículo describe la sexta edición del Taller de Análisis Semántico en la SEPLN, conocido como TASS 2017. TASS tiene como objetivo principal incentivar la investigación y desarrollo de recursos, técnicas, algoritmos y herramientas para tareas relacionadas con el análisis semántico en español. A continuación, se describen las tareas propuestas para la edición 2017, así como los corpus creados y utilizados, los distintos participantes y los resultados obtenidos.
Palabras clave: TASS 2017, análisis de opiniones, análisis semántico

1 Introduction

Since some years ago, Natural Language Processing (NLP) researchers have been working on the discovery of the meaning of utterances from different perspectives. One of those perspectives is the understanding of the subjective information or rather opinion information. The task of Sentiment Analysis (SA) is the result of this study, and it is defined as the computational treatment of opinion, sentiment and subjectivity in text (Pang and Lee, 2008).

However, the potential semantic information encoded in an utterance is so rich and broad that different new NLP tasks have arosen, such as argumentation mining, stance classification, irony detection or the considered tasks in the different editions of our sibling workshop SemEval.[1]

The Spanish language is the second native language in the world and the second language in number of speakers. Nevertheless, the progress of the NLP research in Spanish is far away to the advance of other languages like English. Consequently, TASS[2] (*Taller de Análisis de Sentimientos en la SEPLN* / Workshop on Sentiment Analysis at SEPLN) was born in 2012 with the aim of fostering the development of specific NLP techniques for the computational treatment of opinions of text written in Spanish. The previous editions in 2016 (García-Cumbreras et al., 2016), 2015 (Villena-Román et al., 2015), 2014 (Villena Román et al., 2015), 2013 (Villena-Román et al., 2014) and 2012 (Villena-Román et al., 2013) have yielded outstanding linguistic resources such as the General Corpus of TASS and some datasets for the task of polarity classification at aspect level, used by a great number of research groups and companies as reference for Spanish. Additionally, a research community has been created around TASS that usually participate in the workshop and contribute with vivid discussions about the state-of-the-art and the next challenges in SA in Spanish.

[1]`http://alt.qcri.org/semeval2018/`

[2]`http://www.sepln.org/workshops/tass`

ISSN 1613-0073

The organization committee of the workshop has updated its name in the edition of 2017 because of the need of widening the gamut of semantic tasks in TASS. The new name of TASS is Workshop on Semantic Analysis at SEPLN (*Taller de Análisis Semántico en la SEPLN*), which allows to keep the acronym TASS. The change of the name is a call to researchers on other semantic tasks (argumentation mining, irony detection, stance classification...) to organize a shared-task for the treatment of semantic information in Spanish for the next edition.

The edition of 2017 proposes two subtasks, polarity classification at document (tweet) level (Task 1) and aspect level polarity classification (Task 2). Apart from reusing several datasets of previous editions, a new dataset was specifically generated for this edition. The new dataset is called InterTASS, which is composed of more than 2,000 tweets annotated at four opinion intensity level (POSITIVE, NEUTRAL, NEGATIVE and NONE). Further details about the tasks and the datasets in Sections 2 and 3 respectively.

The edition of 2017 has attracted the participation of 11 teams, mainly from Spain and America. Most of the systems follow the state-of-the-art of SA, which is the use of a deep learning architecture. Most of the teams participated in Task 1, and a few of them in Task 2, which is an indication that polarity classification at aspect level is a tough task.

The rest of this paper is organized as follows. Section 2 presents in more details the two subtasks of TASS 2017. Section 3 describes the datasets and how we created them. Section 4 presents the submitted systems and the results reached by them. Finally, Section 5 concludes and points the future work in TASS.

2 Tasks

TASS 2017 has proposed two tasks addressing the challenging task of SA in Twitter in Spanish.

2.1 Task 1. Sentiment Analysis at Tweet level

This main task focused on the evaluation of polarity classification systems at tweet level in Spanish. Systems were evaluated on three different datasets: the two test sets of the General Corpus of TASS[3] and a new corpus, InterTASS, which was specifically developed in 2017 for the task (see Section 3).

Datasets were annotated with 4 different polarity labels POSITIVE, NEGATIVE, NEUTRAL and NONE), and systems had to identify the intensity of the opinion expressed in each tweet in any of those 4 intensity levels. For the two sets of the General Corpus of TASS, which was annotated in 6 polarity tags, a direct translation from P+ into P and N+ into N was performed so that the evaluation is consistent with InterTASS and based on 4 levels of intensity of polarity.

All datasets were divided into training, development and test datasets, which were provided to participants in order to train and evaluate their systems. Systems were allowed to use any set of data as training dataset, i.e. the training set of InterTASS, other training sets from the previous editions of TASS or other sets of tweets. However, using the test set of InterTASS and the test set of the datasets of previous editions as training data was obviously forbidden. Apart from that, participants could use any kind of linguistic resource for the development of their classification model.

Participants were expected to submit 3 experiments per each evaluation set, so each participant team could submit a maximum of 9 files of results. Results must be submitted in a plain text file with the following format:

```
tweet_id \t polarity
```

Allowed values for polarity were P, NEU, N and NONE.

Accuracy and the macro-averaged versions of Precision, Recall and F1 were used as evaluation measures. Systems were be ranked by the Macro-F1 and Accuracy measures.

2.2 Task 2. Aspect-based Sentiment Analysis

This second task proposed the development of aspect-based polarity classification systems. Two datasets from previous editions were used to evaluate the systems: Social-TV and STOMPOL (see Section 3). The two datasets were annotated for aspect, the main

[3]The entire test set annotated with 4 classes, the 1k test set also annotated with 4 classes.

category of aspect, and the polarity of the opinion about the aspect. Systems had to classify the opinion about the given aspect in 3 different polarity labels (POSITIVE, NEGATIVE, NEUTRAL).

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

```
tweetid \t aspect \t polarity
```

Allowed polarity values were P, NEU and N.

For evaluation, exact match with a single label combining "aspect-polarity" was used. Similarly to Task 1, the macro-averaged version of Precision, Recall and F1, and Accuracy were the evaluation measures, and Macro-F1 were used for ranking the systems.

3 Datasets

TASS 2017 provides four datasets to the participants for the evaluation of their systems. Three of the datasets were used in previous editions, and a new dataset was created for TASS 2017, namely InterTass.

The datasets will be made freely available to the community after the workshop.[4]

3.1 InterTASS

International TASS Corpus (*InterTASS*) is a new corpus released this year for general task (Task 1). The goal of the organization of TASS is the creation of a corpus of tweets written in the Spanish language spoken in Spain and in different Hispano-American countries. We release the first version of InterTASS in TASS 2017, which is only composed of tweets posted in Spain and written in the Spanish language spoken in Spain.

More than 500,000 tweets were collected, from July 2016 to January 2017, using some keywords. The downloaded set of tweets was filtered out in order to meet the following requirements:

- The language of the tweets must be Spanish[5],
- each tweet must contain at least one adjective,
- the minimum length of each tweet must be four words.

Then, the general sentiment of a random selection of tweets was manually annotated by five annotators. We used a scale of 4 levels of polarity: positive (P), neutral (NEU), negative (N) and no sentiment tag (NONE). Each tweet was finally annotated at least by three annotators. When a tweet has the same tag by two of more annotators, the process end. If not, each annotator revised the tweet again, until it has the same tag by two of more annotators. The annotation resulted in a corpus of 3,413 tweets, which was split into three datasets: training, development and test. Table 1 shows the size of each dataset of InterTASS corpus.

Corpus	**#Tweets**
Training	1,008
Developement	506
Test	1,899
Total	3,413

Table 1: Number of tweets in each dataset of InterTASS

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,[6] 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 training set was released, so the participants could train and validate their models. The test corpus was provided without any annotation and has been used to evaluate the results. The InterTass statistics are in Table 2.

	Training	Dev.	Test
P	317	156	642
NEU	133	69	216
N	416	219	767
NONE	138	62	274
Total	1,008	506	1,899

Table 2: Number of tweets in each dataset and class of InterTASS

The three datasets of the corpus are three XML files. Figure 1 shows an example of an InterTASS XML file.

[4]Further information for requesting the datasets in: `http://www.sepln.org/workshops/tass/`.

[5]We used the python library `langdetect`.

[6]`https://dev.twitter.com/terms/api-terms`

```
<tweet>
  <tweetid>768212591105703936</
     tweetid>
  <user>martitarey13</user>
  <content>@estherct209 jajajaja la
     tuya y la d mucha gente seguro
     !! Pero yo no puedo sin mi
     melena me muero </content>
  <date>2016-08-23 22:25:29</date>
  <lang>es</lang>
  <sentiment>
    <polarity>
      <value>N</value>
      <type>AGREEMENT</type>
    </polarity>
  </sentiment>
</tweet>
```

Figure 1: A tweet from the XML file of InterTASS corpus

```
<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>
```

Figure 2: A tweet from the XML file of the General Corpus of TASS

3.2 General corpus

The General Corpus of TASS has 68,000 tweets, written in Spanish by about 150 well-known personalities and celebrities of the world of politics, economy, communication, mass media and culture, between November 2011 and March 2012. The details of the corpus are described in (Villena-Román et al., 2015; García-Cumbreras et al., 2016). Figure 2 shows a tweet from the General Corpus of TASS.

3.3 Social-TV Corpus

The Social-TV corpus was collected during the 2014 Final of Copa del Rey championship in Spain between Real Madrid and F.C. Barcelona, played on 16 April 2014 at Mestalla Stadium in Valencia. After filtering out useless information a subset of 2.773 tweets was selected. The details of the corpus are described in (Villena-Román et al., 2015; García-Cumbreras et al., 2016).

All tweets were manually annotated with 31 different aspects and its sentiment polarity. It was randomly divided into training set (1.773 tweets) and test set (1.000 tweets), with a similar distribution of both aspects and sentiments.

Figure 3 shows a tweet from the Social-TV corpus.

```
<tweet id="456544894501146625">
Para mi, <sentiment aspect="Jugador-
    Isco" polarity="P">ISCO</
    sentiment>
ha hecho un <sentiment aspect="
    Partido" polarity="P">partidazo</
    sentiment>.
<sentiment aspect="Partido" polarity=
    "P">El mejor partido</sentiment>
    desde que llego al
<sentiment aspect="Equipo-Real_Madrid
    " polarity="NEU">Real Madrid</
    sentiment>.
</tweet>
```

Figure 3: A tweet from the XML file of the Social-TV corpus

3.4 STOMPOL

STOMPOL (corpus of Spanish Tweets for Opinion Mining at aspect level about POLitics) is a corpus of Spanish tweets developed for the research in opinion mining at aspect level. 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.

The corpus is composed of 1,284 tweets, and has been divided into training set (784 tweets), which is provided for building and validating the systems, and test set (500 tweets) that will be used for evaluation. The details of the corpus are described in (Villena-Román et al., 2015; García-Cumbreras et al., 2016). Figure 4 shows a tweet from the STOMPOL corpus.

```
<tweet id="591172256971280385">
@rosadiezupyd lamenta que el #
<sentiment aspect="Economia" entity="
    Union_Progreso_y_Democracia"
    polarity="N">empleo</sentiment>
    no termine de estabilizarse y
    dice que el
<sentiment aspect="Economia" entity="
    Union_Progreso_y_Democracia"
    polarity="N">#paro</sentiment> "
    sigue siendo dramÃ¡tico" http://t
    .co/1xdS3UjJWk #EPA
</tweet>
```

Figure 4: STOMPOL XML example

4 Participants and Results

Most of the systems submitted in TASS 2017 are based on the use of deep learning techniques as the state-of-the-art in SA in Twitter. However, some of the systems are based on traditional machine learning methods and others are meta-classifiers whose inputs are the output of deep learning systems and traditional machine learning algorithms. We depict the main features of the systems submitted in the subsequent paragraphs.

Table 3, Table 4 and Table 5 show the results reached by the submitted systems in Task 1, using the test sets of InterTASS corpus and the General Corpus (full test and 1k test). Table 6 and Table 7 shows the results reached by the submitted systems in Task 2, using the test sets of Social-TV corpus and STOMPOL corpus respectively.

Hurtado, Pla, and González (2017) participated in the two tasks. They submitted the same system for both tasks, and the only difference between the tasks lies in the characteristics of the input. The input of the first task is the entire tweet, meanwhile the input in the second task is the context of the aspects, which is previously determined. The authors created a set of domain-specific word embeddings following the approach of Tang (2015). The former word embeddings set is jointly used with a general-domain set of embeddings to represent the tokens of the tweets. The authors evaluated three different neural networks architectures, the first one is a multilinear perceptron (MLP), the second encodes the tweets with a convolutional recurrent neural network (CNN) and the third one with a long-short term memory (LSTM) recurrent neural network (RNN). The performance of each configuration depends on the training set used.

Cerón-Guzmán (2017) presented an ensemble classifier system for the first task. The author generated quantitative features from the tweets, such as the number of words in upper case and the number of words with repeated letters. Moreover, the system used lists of opinion bearing words like iSOL (Molina-González et al., 2013), as well as the inversion of the polarity of words following a window shifting approach for negation handling. The base classifiers of the ensemble system were Logistic Regression and SVM. The system followed two ensemble strategies, namely stacking and maximum classification confidence. The maximum confidence strategy outperformed the stacking strategy and it reached the highest accuracy value with the test set of the InterTASS dataset.

Montañés Salas et al. (2017) used the classifier FastText (Joulin et al., 2016) for only classifying the test set of the InterTASS dataset. The authors performed a traditional pre-processing to the input tweets, however the substitution of words with a emotional meaning by their synonyms from a list of words with a emotional meaning (Bradley and Lang, 1999) stands out.

Rosá et al. (2017) participated in the two tasks. Concerning the first task, the authors submitted three systems: 1) a SVM classifier with word embeddings and quantitative linguistic properties as features; 2) a deep neural network grounded on the use of a CNN for encoding the input tweets; and 3) the combination of the two previous classifiers by the selection of the output class with a higher probability mean from the two previous classifiers. The third strategy outperformed the other ones in two test sets of Task 1. Regarding the Task 2, the authors submitted two SVM classifiers mainly grounded on the use word embeddings.

García-Vega et al. (2017) submitted four systems for the classification of the test set of the InterTASS dataset. The first two systems are a SVM classifier that uses word-embeddings as features. The difference between these two systems lies in the use of additional tweets from the users of the training set. The intention of the authors was the in-

System	M-F1	Acc.
ELiRF-UPV-run1	0.493	0.607
RETUYT-svm_cnn	0.471	0.596
ELiRF-UPV-run3	0.466	0.597
ITAINNOVA-model4	0.461	0.576
jacerong-run-2	0.460	0.602
jacerong-run-1	0.459	0.608
INGEOTEC-evodag_001	0.457	0.507
RETUYT-svm	0.457	0.583
tecnolengua-sent_only	0.456	0.582
ELiRF-UPV-run2	0.450	0.436
ITAINNOVA-model3	0.445	0.561
RETUYT-cnn3	0.443	0.558
SINAI-w2v-nouser	0.442	0.575
tecnolengua-run3	0.441	0.576
tecnolengua-sent_only_fixed	0.441	0.595
ITAINNOVA-model2	0.436	0.576
LexFAR-run3	0.432	0.541
LexFAR-run1	0.430	0.539
jacerong-run-3	0.430	0.576
SINAI-w2v-user	0.428	0.569
INGEOTEC-evodag_002	0.403	0.515
OEG-victor2	0.395	0.451
OEG-victor0	0.383	0.433
OEG-laOEG	0.377	0.505
LexFAR-run2	0.372	0.490
GSI-sent64-189	0.371	0.524
SINAI-embed-rnn2	0.333	0.391
GSI-sent64-149-ant-2	0.306	0.479
GSI-sent64-149-ant	0.000	0.000

Table 3: Task 1 InterTASS corpus results

troduction of the use of language of each user in the classification. The two last systems are deep neural networks grounded on the use of LSTM RNN for the encoding of the meaning of the input tweets. The first neural architecture uses word embeddings as features, and the second one the TF-IDF value of each word of the tweets.

Moctezuma et al. (2017) participation was based on an ensemble of SVM classifiers combined into a non-linear model created with genetic programming to tackle the task of global polarity classification at tweet level. They used B4MSA algorithm, a proposed entropy-based term weighting scheme, which is a baseline supervised learning system based on the SVM classifier, an entropy-based term-weighting scheme. Additionally they used EvoDAG, a GP system that combines all decision values predicted by B4MSA systems. They also used two external datasets to train the B4MSA algorithm.

System	M-F1	Acc.
INGEOTEC-evodag_003	0.577	0.645
jacerong-run-1	0.569	0.706
jacerong-tass_2016-run_3	0.568	0.705
ELiRF-UPV-run2	0.549	0.659
ELiRF-UPV-run3	0.548	0.725
RETUYT-svm_cnn	0.546	0.674
jacerong-run-2	0.545	0.701
ELiRF-UPV-run1	0.542	0.666
RETUYT-cnn	0.541	0.638
RETUYT-cnn3	0.539	0.654
tecnolengua-run3	0.528	0.657
tecnolengua-final	0.517	0.632
tecnolengua-531F1_no_ngrams	0.508	0.652
INGEOTEC-evodag_001	0.447	0.514
OEG-victor2	0.389	0.496
INGEOTEC-evodag_002	0.364	0.449
OEG-laOEG	0.346	0.407
GSI-64sent99ally	0.324	0.434

Table 4: Task 1 General Corpus of TASS (full test) results

Navas-Loro and Rodríguez-Doncel (2017) participated only on Task 1. They experimented with two classifier algorithms, Multinominal Naïve Bayes and Sequential Minimal Optimization for SVM. Furthermore they used morphosyntactic analyses for negation detection, along with the use of lexicons and dedicated preprocessing techniques for detecting and correcting frequent errors and expressions in tweets.

Araque et al. (2017) have proposed, for Task 1, a RNN architecture composed of LSTM cells followed by a feed-forward network. The architecture makes use of two different types of features: word embeddings and sentiment lexicon values. The recurrent architecture allows them to process text sequences of different lengths, while the lexicon inserts directly into the system sentiment information. Two variations of this architecture were used: a LSTM that iterates over the input word vectors, and on the other

System	M-F1	Acc.
RETUYT-svm	0.562	0.700
RETUYT-cnn4	0.557	0.694
RETUYT-cnn2	0.555	0.694
INGEOTEC-evodag_003	0.526	0.595
tecnolengua-run3	0.521	0.638
ELiRF-UPV-run1	0.519	0.630
jacerong-tass_2016-run_3	0.518	0.625
jacerong-run-1	0.508	0.678
jacerong-run-2	0.506	0.673
ELiRF-UPV-run2	0.504	0.596
tecnolengua-final	0.488	0.618
tecnolengua-run4	0.483	0.612
ELiRF-UPV-run3	0.477	0.588
INGEOTEC-evodag_002	0.439	0.431
INGEOTEC-evodag_001	0.388	0.486
OEG-victor3b	0.367	0.386
OEG-victor2	0.366	0.412
OEG-laOEG	0.346	0.448
GSI-run-1	0.327	0.558
GSI-64sent99ally	0.321	0.499

Table 5: Task 1 General Corpus of TASS (1k) results

System	M-F1	Acc.
ELiRF-UPV-run3	0.537	0.615
ELiRF-UPV-run2	0.513	0.600
ELiRF-UPV-run1	0.476	0.625
RETUYT-svm2	0.426	0.595
RETUYT-svm	0.413	0.493

Table 6: Task 2 Social-TV corpus results

hand a combination of the input word vectors and polarity values from a sentiment lexicon.

Tume Fiestas and Sobrevilla Cabezudo (2017) have proposed, for Task 2, an approach based on word embeddings for polarity classification at aspect-level. They used word embeddings to get the similarity between words selected from a training set and make a model to classify each polarity of each aspect for each tweet. Their results show that the more tweets are used, the better accuracy is obtained.

Reyes-Ortiz et al. (2017) have proposed, for Task 1, a system that uses machine learning, vector support machines algorithm and lexicons of semantic polarities at the level of lemma for Spanish. Features extracted from lexicons are represented by the bag-of-word model and they are weighted using Term Frequency measure at tweet level.

System	M-F1	Acc.
ELiRF-UPV-run1	0.537	0.615
RETUYT-svm2	0.508	0.590
ELiRF-UPV-run3	0.486	0.578
ELiRF-UPV-run2	0.486	0.541
C100T-PUCP-run3	0.445	0.528
C100T-PUCP-run1	0.415	0.563
C100T-PUCP-run2	0.414	0.517
RETUYT-svm	0.377	0.514

Table 7: Task 2 STOMPOL corpus results

Moreno-Ortiz and Pérez Hernández (2017) have proposed, for Task 1, a classification model based on the Lingmotif Spanish lexicon, and combined this with a number of formal text features, both general and CMC-specific, as well as single-word keywords and n-gram keywords. They use logistic regression classifier trained with the optimal set of features, SVM classifier on the same features set. Sentiment features are obtained with Lingmotif SA engine (sentiment feature set, text feature set and keywords feature set).

5 *Conclusion and Future work*

TASS was the first workshop about sentiment analysis focused on the processing of texts written in Spanish. In this edition, 11 teams participated with a total of 123 runs, most of them in the InterTASS task.

Anyway, the released corpora and the reports from participants will for sure be helpful for other research groups approaching these tasks.

The future work will mainly go in two directions. On the one hand, the organization of one o more shared-tasks for the treatment of semantic information in Spanish like those mentioned above (argumentation mining, irony detection and stance classification). On the other hand, the extension and improvement of the InterTASS corpus. This corpus has been received with great interest, almost 90% of the experiments have been developed in the first task, so an exhaustive analysis of the behavior of the corpus in this task will shows the right way for a new version of the corpus.

TASS corpora will be released after the workshop for free use by the research community.

Acknowledgement

This research work is partially supported by the project REDES (TIN2015-65136-C2-1-R) and a grant from the Fondo Europeo de Desarrollo Regional (FEDER).

References

Araque, O., R. Barbado, J. F. Sánchez-Rada, and C. A. Iglesias. 2017. Applying recurrent neural networks to sentiment analysis of spanish tweets. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*, volume 1896 of *CEUR Workshop Proceedings*, Murcia, Spain, September. CEUR-WS.

Bradley, M. M. and P. J. Lang. 1999. Affective norms for english words (anew): Stimuli, instruction manual, and affective ratings. Technical report, Center for Research in Psychophysiology, University of Florida.

Cerón-Guzmán, J. A. 2017. Classier ensembles that push the state-of-the-art in sentiment analysis of spanish tweets. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*.

García-Cumbreras, M. A., J. Villena-Román, E. Martínez-Cámara, M. C. Díaz-Galiano, M. T. Martín-Valdivia, and L. A. Ureña López. 2016. Overview of tass 2016. In *TASS 2016: Workshop on Sentiment Analysis at SEPLN*, pages 13–21.

García-Vega, M., A. Montejo-Ráez, M. C. Díaz-Galiano, and S. M. Jiménez-Zafra. 2017. Sinai en tass 2017: Clasificación de la polaridad de tweets integrando información de usuario. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*.

Hurtado, L.-F., F. Pla, and J.-A. González. 2017. Elirf-upv en tass 2017: Análisis de sentimientos en twitter basado en aprendizaje profundo. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*.

Joulin, A., E. Grave, P. Bojanowski, and T. Mikolov. 2016. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Moctezuma, D., M. Graff, S. Miranda-Jiménez, E. S. Tellez, A. Coronado, C. N. Sánchez, and J. Ortiz-Bejar. 2017. A genetic programming approach to sentiment analysis for twitter: Tass'17. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*, volume 1896 of *CEUR Workshop Proceedings*, Murcia, Spain, September. CEUR-WS.

Molina-González, M. D., E. Martínez-Cámara, M.-T. Martí-Valdivia, and J. M. Perea-Ortega. 2013. Semantic orientation for polarity classification in spanish reviews. *Expert Systems with Applications*, 40(18):7250 – 7257.

Montañés Salas, R. M., R. del Hoyo Alonso, J. Vea-Murguía Merck, R. Aznar Gimeno, and F. J. Lacueva-Pérez. 2017. FastText como alternativa a la utilización de deep learning en corpus pequeños. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*.

Moreno-Ortiz, A. and C. Pérez Hernández. 2017. Tecnolengua lingmotif at tass 2017: Spanish twitter dataset classification combining wide-coverage lexical resources and text features. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*, volume 1896 of *CEUR Workshop Proceedings*, Murcia, Spain, September. CEUR-WS.

Navas-Loro, M. and V. Rodríguez-Doncel. 2017. Oeg at tass 2017: Spanish sentiment analysis of tweets at document level. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*, volume 1896 of *CEUR Workshop Proceedings*, Murcia, Spain, September. CEUR-WS.

Pang, B. and L. Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135.

Reyes-Ortiz, J. A., F. Paniagua-Reyes, B. Priego-Sánchez, and M. Tovar. 2017. Lexfar en la competencia tass 2017: Análisis de sentimientos en twitter basado en lexicones. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*, volume 1896 of *CEUR Workshop Proceedings*, Murcia, Spain, September. CEUR-WS.

Rosá, A., L. Chiruzzo, M. Etcheverry, and S. Castro. 2017. Retuyt en tass 2017: Análisis de sentimientos de tweets en español utilizando svm y cnn. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*.

Tang, D. 2015. Sentiment-specific representation learning for document-level sentiment analysis. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, WSDM '15, pages 447–452, New York, NY, USA. ACM.

Tume Fiestas, F. and M. A. Sobrevilla Cabezudo. 2017. C100tpucp at tass 2017: Word embedding experiments for aspect-based sentiment analysis in spanish tweets. In *Proceedings of TASS 2017: Workshop on Sentiment Analysis at SEPLN co-located with 33nd SEPLN Conference (SEPLN 2017)*, volume 1896 of *CEUR Workshop Proceedings*, Murcia, Spain, September. CEUR-WS.

Villena-Román, J., J. García-Morera, M. A. García-Cumbreras, E. Martínez-Cámara, M. T. Martín-Valdivia, and L. A. Ureña López. 2015. Overview of tass 2015. In *TASS 2015: Workshop on Sentiment Analysis at SEPLN*, pages 13–21.

Villena-Román, J., J. García-Morera, S. Lana-Serrano, and J. C. González-Cristóbal. 2014. Tass 2013 - a second step in reputation analysis in spanish. *Procesamiento del Lenguaje Natural*, 52(0):37–44, March.

Villena-Román, J., S. Lana-Serrano, E. Martínez-Cámara, and J. C. González-Cristóbal. 2013. Tass - workshop on sentiment analysis at sepln. *Procesamiento del Lenguaje Natural*, 50:37–44.

Villena Román, J., E. Martínez Cámara, J. García Morera, and S. M. Jiménez Zafra. 2015. Tass 2014 - the challenge of aspect-based sentiment analysis. *Procesamiento del Lenguaje Natural*, 54(0):61–68.