

# Towards the Improvement of Automatic Emotion Pre-annotation with Polarity and Subjective Information

**Lea Canales**

University of Alicante  
Alicante (Spain)  
lcanales@dlsi.ua.es

**Walter Daelemans**

University of Antwerp  
Antwerpen (Belgium)  
walter.daelemans@uantwerpen.be

**Ester Boldrini**

University of Alicante  
Alicante (Spain)  
eboldrini@dlsi.ua.es

**Patricio Martínez-Barco**

University of Alicante  
Alicante (Spain)  
patricio@dlsi.ua.es

## Abstract

Emotion detection has a high potential positive impact on the benefit of business, society, politics or education. Given this, the main objective of our research is to contribute to the resolution of one of the most important challenges in textual emotion detection: emotional corpora annotation. This will be tackled by proposing a semi-automatic methodology. It consists in two main phases: (1) an automatic process to pre-annotate the unlabelled sentences with a reduced number of emotional categories; and (2) a manual process of refinement where human annotators will determine which is the dominant emotion between the pre-defined set. Our objective in this paper is to show the pre-annotation process, as well as to evaluate the usability of subjective and polarity information in this process. The evaluation performed confirms clearly the benefits of employing the polarity and subjective information on emotion detection and thus endorses the relevance of our approach.

## 1 Introduction

The creation of a labelled emotion corpus is not trivial; detecting emotion in text can be difficult even for humans due to the influence of each one's own background that affects emotion interpretation.

Most relevant research carried out so far has shown the difficulties related to this task, such as obtaining a good inter-annotator agreement (IAA) or the time required for its development. This is due to the fact that manual annotations can be significantly influenced by a set of different factors,

such as clarity of instructions, the difficulty of the task, and even by the annotation scheme (Mohammad, 2016). For this reason, in this paper, a semi-automatic methodology is proposed with the aim of achieving a highly reliable of this task, reducing its complexity automatically.

The methodology proposed in our research consists of two main phases: (1) an automatic process to pre-annotate the unlabelled sentences with a reduced number of emotional categories; and (2) a manual process of refinement, where human annotators will determine which is the dominant emotion between the emotional categories selected in phase 1.

This means that if we want to annotate the data with the Ekman's emotions (Ekman, 1992) (ANGER, DISGUST, FEAR, JOY, SADNESS and SURPRISE) plus the NEUTRAL category, our methodology would reduce automatically the number of emotional categories proposed to the annotators from seven to three. For instance, given the sentence *I've never missed anyone so much as you*, the pre-annotation process selects SADNESS, FEAR and DISGUST as possible dominant emotions of this sentence. In the manual process of refinement (the second phase), the annotators determine between these three emotions which one is the best to represent the dominant emotion.

As an annotation methodology, our aim is to reduce the number of emotional categories automatically, since the number of coding categories influences reliability estimation. As Antoine et al. (2014) concluded, annotation agreement increases significantly when the number of classes decreases. Hence, our hypothesis is that the decrease of the complexity of emotion annotation through the reduction of the number of emotional categories will allow us to improve its reliability. This methodology would allow us to an-

notate a large amount of emotional data efficiently and with a guarantee of standards of reliability, as has been shown in other NLP tasks (Rehbein et al., 2009; Fort and Sagot, 2010). Our main objective in this paper is to evaluate the usability of subjective and polarity information in the pre-annotation process, as well as showing the method proposed in detail.

The proposed pre-annotation process is based on the use of distributed representations of the emotion words and the sentences that we want to annotate. A big advantage in using these representations that encode semantic information is that they can be generated from large corpora of unlabelled text, and can be trained on very large corpora in a reasonable amount of time. Thus, it is simple to filter the number of emotional categories that can be associated with each sentence and reduce the ambiguity of the second phase of this methodology that will determine the dominant emotion.

The evaluation is carried out with the Aman corpus (Aman and Szpakowicz, 2007), one of the most relevant corpora at present employed in textual emotion recognition, and the results reflect clearly the benefits of employing the subjective and polarity information in the pre-annotation process.

The rest of the paper is organized as follows. Section 2 presents related work. After this, the proposed method is described in detail in Section 3. Then, in Section 4 the test corpus and the word embeddings employed are presented. Section 5 is aimed at showing the evaluation methodology, the results obtained and a discussion of these results. Finally, Section 6 details our conclusions and future work.

## 2 Related Work

Training a supervised machine-learning based textual emotion recognition system implies manually annotating a significant amount of text.

With the aim of training emotional machine-learning approaches of different domains and from a variety of genres, several manual-annotated corpora have been developed, such as the 185 children stories annotated with emotion categories (Alm et al., 2005); blog posts collected directly from Web with emotion categories and intensity (Aman and Szpakowicz, 2007); or a news headlines corpus with emotion categories and valence

(Strapparava and Mihalcea, 2007).

The cost of this, in terms of human effort, slows down the development of an accurate emotion recognition system. This problem is shared by other NLP tasks and one usual way to improve this situation is to automatically pre-annotate the corpus so that the work of the annotators is limited to the validation of this pre-annotation.

Pre-annotation has been widely studied in NLP tasks, such as Named Entity Recognition (NER), Part Of Speech (POS) tagging and Semantic Frame/Role Labelling, reporting a gain in time and quality.

Marcus et al. (1993) is one of the first approaches where the pre-annotation process is assessed for POS tagging. In this work, the model of annotation consists of two stages: 1) automatic POS assignment and 2) manual correction is evaluated to determine how to maximise the speed, inter-annotator consistency, and accuracy of POS tagging. The experiment showed that manual tagging took about twice as long as correcting, with about twice the inter-annotator disagreement rate and an error rate that was about 50% higher. More recently, Fort and Sagot (2010) evaluate the influence of automatic pre-annotation on the manual POS annotation of a corpus, both from the quality and the time points of views, with specific attention to biases. Their experiments confirmed and detailed the gain in quality and demonstrated that even a not so accurate tagger can help improve annotation speed.

Rehbein et al. (2009) led quite thorough experiments on the subject, in the field of semantic frame assignment annotation. Although in this case, the results of the experiments are a bit disappointing as they could not find a direct improvement of annotation time using pre-annotation, they found that noisy and low-quality pre-annotation does not overall corrupt human judgement.

More recently, Lingren et al. (2014) evaluate the impact of pre-annotation on annotation speed and potential bias for clinical named entity recognition in clinical trial announcements. As in other NLP tasks, they concluded that the annotator with the pre-annotated text needed less time to annotate than the annotator with non-labeled text. Moreover, the pre-annotation did not reduce the IAA or annotator performance.

Consequently and with the aim of overcoming the cost and time-consuming shortcoming of man-

ual annotation, the objective of this research is to propose a new semi-automatic methodology for large-scale annotation of emotional corpora and with standards of reliability.

### 3 Pre-annotation Process

This section describes the pre-annotation process developed to improve the emotion annotation task. The section is divided into two subsections where the main tasks carried out by the process are explained in detail.

The process receives as input data a collection of unlabelled sentences/phrases and a set of emotions. The approach presented in this paper works with Ekman's six basic emotions (Ekman, 1992) since it is the set of emotional categories employed in the test corpus and is one of the emotional model most employed in textual emotion detection. Although the process can also be adapted for another group of emotions.

The pre-annotation process consists of two main steps: the representation of emotional categories and sentences in a semantic space, a vector space model of meaning, and the association between emotions and sentences, explained in subsection 3.1 and subsection 3.2, respectively.

#### 3.1 Emotional Categories and Sentences in Semantic Space

The first step towards data annotation consists in encoding the emotions and the sentences in a semantic space with the help of distributed representations. This step is split into two main sub-steps: pre-processing data and the creation of the distributional vectors shown in Figure 1.

With respect to the pre-processing in emotions, it consists in creating a bag of words (a seed) related to each emotion by exploring an emotional lexicon. The process selects the words associated with only one of the Ekman's basic emotions to create an accurate seed without ambiguous words. In this approach, EmoSenticNet is employed as emotion lexicon.

EmoSenticNet (Poria et al., 2013) is a lexical resource of 13,171 words that assigns qualitative emotion labels (Ekman's emotions: *anger*, *fear*, *disgust*, *sadness*, *surprise*, or *joy*) and quantitative polarity scores to SenticNet concepts (Cambria et al., 2016). As the process works with the words associated with only one of the emotion, the resource is reduced to 2,289 words.

About the pre-processing of the sentences, it is carried out by tokenizing each sentence.

The second step of the representation in semantic space is common for emotions and sentences and consists in creating emotion vectors and sentence vectors by replacing each word in every bag of words with the vector representation of such word. This creates an  $M \times n$  matrix for each emotion and sentence, where  $M$  is the dimensionality of the word vectors and  $n$  the number of words contained in the bag of words. Following this, for each emotion and sentence, we then obtain a single vector of dimensionality  $M$  by applying averaging as a compositional function. Whereby obtaining so-called *emotion vectors* and *sentence vectors* which represent the entire meaning of each emotion and sentence in one vector, respectively.

#### 3.2 Associating Sentences with Emotions

Because all emotions and sentences are created using the same distributed vectors and compositional function, the vector space in which they are placed is also comparable. Hence, we use the cosine distance between emotions and sentences to associate them. Concretely for each sentence, the process measures the similarity and those three emotions whose semantic similarity is higher are associated (*original approach*). The process pre-annotates with three emotions because it is half of the number of Ekman's basic emotions. If the process works with a greater or less group of emotion categories, the number of emotions pre-annotated will be increased or reduced respectively.

In addition to this approach, the usability of the polarity and subjective information is also evaluated (*polarity-subjectivity approach*). To achieve that, the order of the emotions proposed by the system is re-ordered according to the polarity and subjective values of each sentence. For that, the Sentiment Analysis tool from Pattern (De Smedt and Daelemans, 2012) is employed, which returns an averaged (polarity, subjectivity) tuple for a given string.

About the classification of the Ekman's six basic emotions according to the polarity, we assume that JOY belongs to the positive class, while the other five emotions have negative polarity, except for SURPRISE since it can be employed from the positive and negative point of view. Hence, when SURPRISE is the first emotion proposed by the system and the subjective value is not zero, the po-

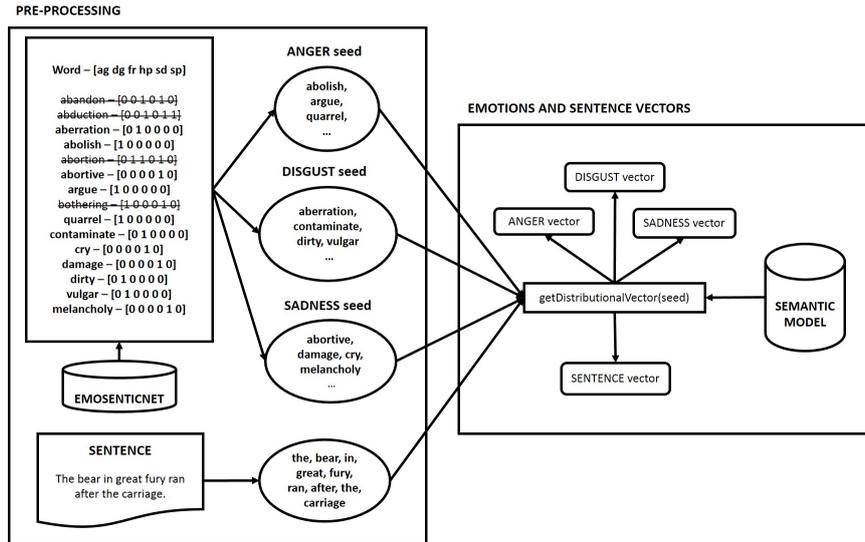


Figure 1: The process of representing the emotions and sentences in the semantic space.

larity information is employed to re-order the rest of the emotions.

The re-ordering of the emotions is carried out considering the following conditions:

- If the subjective value is zero, the sentence will be considered NEUTRAL and thus this category is proposed in the first place. In another case, the polarity value will evaluate it.
- If the polarity value is POSITIVE (higher than zero), the emotion considered positive (JOY) is proposed in the first position.
- If the polarity value is NEGATIVE (less than zero), the emotions considered negatives (ANGER, DISGUST, FEAR, SADNESS) are proposed before the positives ones. The order between these emotions is determined by the semantic similarity obtained when *emotion word vectors* are compared to *sentence vector*.

Table 1 shows examples of how the polarity and subjectivity information is employed in the pre-annotation process.

## 4 Materials

### 4.1 Data

The Aman corpus (Aman and Szpakowicz, 2007) is the dataset employed to evaluate the approaches. It contains sentence-level annotation of 4,000 sentences from blogs posts collected directly from

Web. This corpus was annotated manually with the six emotion categories proposed by Ekman plus the emotion intensity (high, medium, or low).

The reasons for choosing this corpus for testing the approach are: (i) this corpus has been employed in many emotion studies as a benchmark (Keshtkar and Inkpen, 2010) (Chaffar and Inkpen, 2011) (Mohammad, 2012); and (ii) it is possible to check the effectiveness of our approach in one of the Social Media platforms: the blog post. A platform where emotion detection would be helpful since this genre allows people to post messages to share information and opinions.

### 4.2 Word Vectors

Given that the pre-annotation process is based on distributional representations, different Distributional Semantic Models (DSM) have been evaluated. All of them have been employed in the two approaches carried out in this research: the *original approach* and the *polarity-subjective approach*. Concretely, our approaches have been evaluated using four semantic spaces.

*Vector Space Model (baseline)*: A simple semantic space is built by a Vector Space Model (VSM) created with EmoSenticNet, the emotional lexicon. In this space, the emotions and sentences are represented by a vector that contains information about which EmoSenticNet words occur in each sentence or emotion.

*Affective Space* (Cambria et al., 2015): this set is the 100-dimensional vector space representation of AffectNet (a matrix of affective commonsense

Sentence	1st ranking	Polarity	Subjectivity	Emotions proposed
This was the best summer I have ever experienced.	joy, disgust, sadness, fear, surprise, anger	0.9	0.6	joy, disgust, sadness
I hate fucking pills.	anger, surprise, fear, disgust, sadness, joy	-0.7	0.85	anger, surprise, fear
Had a lovely birthday yesterday with Alex and Christine.	sadness, joy, disgust, surprise, fear, anger	0.5	0.75	joy, sadness, disgust
I'm becoming a broken toy and now that I have had twelve (I counted) vials of blood drawn, I just feel like I'm completely useless.	joy, sadness, disgust, fear, surprise, anger	-0.15	0.48	sadness, disgust, fear
You don't know their middle name or the age of their sister.	joy, disgust, sadness, fear, surprise, anger	0.0	0.0	neutral, joy, disgust

Table 1: Examples of pre-annotation process. The *1st ranking* column shows the order proposed by the system before employing the polarity and subjective information. The *Emotion proposed* column shows the pre-annotated emotions by the system after re-ordering the first ranking.

knowledge in which common-sense concepts are linked to semantic and affective features).

*Common Crawl GloVe vectors* (Pennington et al., 2014): this set contains the 300-dimension GloVe vectors trained on 42 billion tokens of web data from Common Crawl<sup>1</sup>, a data collection over the last 7 years that contains raw web page data, extracted metadata and text extractions.

*Ultradense Sentiment Analysis Word Embeddings* (Rothe et al., 2016): these pre-trained embeddings are the results of learning an orthogonal transformation of the embedding space that focuses the information relevant for a task. In our approach, the 300-dimension Google News vectors focused on Sentiment Analysis in dimension 1 are employed.

## 5 Evaluation

It is necessary to develop an internal and external evaluation to assess the semi-automatic methodology on emotional data. The internal evaluation involves assessing the pre-annotation process whereas, the external evaluation has as objective the evaluation of the work of the annotators in the second phase of the methodology. In this paper, the internal evaluation is carried out.

In the internal evaluation, the pre-annotation process is assessed measuring the precision (P), recall (R), and F1-values (F1) of the emotions proposed by our system against the gold standard of Aman corpus. As the process pre-annotates three emotions, if the correct emotion (the gold standard) is one of these three emotions, the prediction will be considered as correct.

<sup>1</sup><http://commoncrawl.org/>

## 5.1 Results

The results of the experiments on the *original approach* are presented in Table 2, which shows that considering the F1-macro, all models outperform significantly the baseline. The improvements in *Affective Space* are due mainly to their results in DISGUST, FEAR and SADNESS emotions where obtains recall values higher than 50% and high precision values in DISGUST and SADNESS. In the case of the *Common Crawl GloVe* model, the recall values are higher than 90% in JOY, SADNESS and SURPRISE emotions. From these high values, we may draw that these emotions are frequently found between the emotions proposed by the system. About the *Ultradense SA* model, it is surprising the recall obtained in emotions like ANGER, JOY, SADNESS, and SURPRISE and the precision obtained by ANGER and DISGUST, two emotions difficult to detect. The high recall in *Ultradense SA* and *Common Crawl Glove* could be due to the fact that the vocabulary in these models is much larger than the rest of the models.

The results of the performance of the *polarity-subjective approach* are shown in Table 3. In this case, all models also outperform the baseline. Although, the best result is obtained by *Common Crawl Glove* model obtaining the best value for SURPRISE emotion and values close to the other models in the rest of the emotions. Moreover, the results show that the recall obtained by *Common Crawl Glove* and *Ultradense SA* remains high in JOY, SADNESS and SURPRISE and the use of polarity and subjective information allow to improve the recall and precision in ANGER, DISGUST and FEAR emotions. About *Affective*

<i>Original Approach</i>												
	Baseline			Affective Space			Common Crawl GloVe			Ultradense SA		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Anger	0.16	0.08	0.11	0.13	0.08	0.10	0.31	0.04	0.08	0.62	0.23	<b>0.33</b>
Disgust	0.21	0.12	0.15	0.48	0.64	<b>0.55</b>	0.18	0.09	0.12	0.55	0.27	0.37
Fear	0.46	0.17	0.24	0.34	0.50	<b>0.41</b>	0.62	0.04	0.08	0.53	0.07	0.12
Joy	0.16	0.75	0.26	0.15	0.68	0.24	0.15	0.90	0.25	0.23	0.93	<b>0.37</b>
Sadness	0.08	0.14	0.11	0.43	0.54	0.48	0.47	0.95	<b>0.63</b>	0.12	0.95	0.21
Surprise	0.07	0.15	0.09	0.07	0.09	0.08	0.71	0.93	<b>0.80</b>	0.51	0.77	0.61
Neutral	0.69	0.18	0.28	0.75	0.20	<b>0.31</b>	0.75	0.06	0.12	0.69	0.05	0.09
Micro Avg.	0.24	0.24	0.24	0.29	0.29	<b>0.29</b>	0.23	0.23	0.23	0.24	0.24	0.24
Macro Avg.	0.26	0.23	0.18	0.34	0.39	<b>0.31</b>	0.46	0.43	0.30	0.47	0.47	0.30

Table 2: Precision, Recall and F1-values obtained in the *original approach* using different distributional representations.

<i>Polarity-Subjectivity Approach</i>												
	Baseline			Affective Space			Common Crawl GloVe			Ultradense SA		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Anger	0.20	0.08	0.11	0.27	0.07	0.11	0.92	0.25	0.39	0.91	0.44	<b>0.59</b>
Disgust	0.24	0.10	0.15	0.25	0.61	0.36	0.86	0.17	0.29	0.91	0.37	<b>0.53</b>
Fear	0.58	0.13	0.21	0.54	0.61	<b>0.57</b>	1.00	0.33	0.50	0.68	0.11	0.19
Joy	0.23	0.78	0.36	0.28	0.75	0.41	0.30	0.87	0.44	0.32	0.88	<b>0.47</b>
Sadness	0.13	0.14	0.14	0.36	0.52	<b>0.43</b>	0.22	0.95	0.35	0.20	0.88	0.33
Surprise	0.12	0.17	0.14	0.14	0.11	0.12	0.88	0.76	<b>0.81</b>	0.49	0.64	0.55
Neutral	0.82	0.51	<b>0.63</b>	0.84	0.51	<b>0.63</b>	0.89	0.49	<b>0.63</b>	0.89	0.49	<b>0.63</b>
Micro Avg.	0.48	0.48	0.48	0.52	0.52	0.52	0.54	0.54	<b>0.54</b>	0.54	0.54	<b>0.54</b>
Macro Avg.	0.33	0.27	0.25	0.38	0.45	0.38	0.72	0.55	<b>0.49</b>	0.63	0.54	0.47

Table 3: Precision, Recall and F1-values obtained in the *polarity-subjectivity approach* using different distributional representations.

*Space*, the best values are obtained in FEAR and SADNESS where the precision and recall are really interesting.

Comparing the *original* and *polarity-subjective* approaches, it is safe to say that the results clearly reflect the improvements obtained by the polarity-subjective information. Mostly, these enhancements are observed in *Common Crawl GloVe* and *Ultradense SA* models outperforming the F1-value in more than 17% and improving the values obtained by ANGER, DISGUST and FEAR emotions. Thus, the results demonstrate clearly the usability of polarity and subjective information in the pre-annotation process since all models improve its performance when this information is employed.

## 6 Conclusion

In this paper, we provide a new methodology for improving the emotional annotation task. Concretely, the pre-annotation process is presented

and evaluated with the aim to analyzing the usability of polarity and subjective information in this method. Furthermore, to compare the benefits of the *word embeddings* in the process, three models built in very different ways are compared.

The evaluation performed confirms clearly the benefits of employing the polarity and subjective information for emotion detection. Therefore, this will be the approach employed in the next steps of our future work.

Regarding the DSM models, the results show the need of building word embeddings from a large amount of data to obtain a large vocabulary and with high dimensionality to codify more semantic features since the better performance have been achieved by *Common Crawl GloVe* and *Ultradense SA*.

Taking into account the results obtained, our future work will be focused on exploring other *word embeddings* with these features and developing the external evaluation through a manual annotation

task with the sentences pre-annotated by our automatic process.

## Acknowledgments

This research has been supported by the FPI grant (BES-2013-065950) and the research stay grants (EEBB-I-17-12578) from the Spanish Ministry of Science and Innovation. It has also funded by the Spanish Government (ref. TIN2015-65136-C02-2-R and ref. TIN2015-65100-R), the Valencian Government (grant no. PROMETEOII/ 2014/001), the University of Alicante (ref. GRE16-01) and BBVA Foundation (ASAP project).

## References

- Cecilia Ovesdotter Alm, Dan Roth, and Richard Sprout. 2005. Emotions from text: Machine learning for text-based emotion prediction. In *Proc. of the conference on HLT-EMNLP*. pages 579–586.
- Saima Aman and Stan Szpakowicz. 2007. Identifying Expressions of Emotion in Text. In *Text, Speech and Dialogue*, pages 196–205.
- Jean-yves Antoine, Jeanne Villaneau, and Anaïs Lefevre. 2014. Weighted Krippendorff’s alpha is a more reliable metrics for multi-coders ordinal annotations: experimental studies on emotion, opinion and coreference annotation. In *Proc. of the 14th Conference of the European Chapter of the Association for Computational Linguistics*. 1, pages 550–559.
- Erik Cambria, Jie Fu, Federica Bisio, and Soujanya Poria. 2015. Affectivespace 2: Enabling affective intuition for concept-level sentiment analysis. In *Proc. of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA*. pages 508–514.
- Erik Cambria, Soujanya Poria, and Rajiv Bajpai. 2016. SenticNet 4 : A Semantic Resource for Sentiment Analysis Based on Conceptual Primitives. *Sentic.Net*.
- Soumaya Chaffar and Diana Inkpen. 2011. Using a Heterogeneous Dataset for Emotion Analysis in Text. In *Proc. of the 24th Canadian Conference on Advances in Artificial Intelligence*. Springer-Verlag, Berlin, Heidelberg, Canadian AI’11, pages 62–67.
- Tom De Smedt and Walter Daelemans. 2012. Pattern for python. *J. Mach. Learn. Res.* 13(1):2063–2067.
- Paul Ekman. 1992. An argument for basic emotions. *Cognition and Emotion* pages 169–200.
- Karèn Fort and Benoît Sagot. 2010. Influence of pre-annotation on pos-tagged corpus development. In *Proc. of the Fourth Linguistic Annotation Workshop*. Association for Computational Linguistics, Stroudsburg, PA, USA, LAW IV ’10, pages 56–63.
- Fazel Keshtkar and Diana Inkpen. 2010. A Corpus-based Method for Extracting Paraphrases of Emotion Terms. In *Proc. of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. Association for Computational Linguistics, Stroudsburg, PA, USA, CAAGET ’10, pages 35–44.
- Todd Lingren, Louise Deleger, Katalin Molnar, Haijun Zhai, Jareen Meitzen-Derr, Megan Kaiser, Laura Stoutenborough, Qi Li, and Imre Solti. 2014. Evaluating the impact of pre-annotation on annotation speed and potential bias: natural language processing gold standard development for clinical named entity recognition in clinical trial announcements. *Journal of the American Medical Informatics Association : JAMIA* 21(3):406–413.
- Mitchell P. Marcus, Mary Ann Marcinkiewicz, and Beatrice Santorini. 1993. Building a large annotated corpus of english: The penn treebank. *Comput. Linguist.* 19(2):313–330.
- Saif Mohammad. 2012. Portable Features for Classifying Emotional Text. In *Proc. of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Montréal, Canada, pages 587–591.
- Saif M Mohammad. 2016. Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text. In Herb Meiselman, editor, *Emotion Measurement*, Elsevier.
- Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. GloVe: Global Vectors for Word Representation. In *Empirical Methods in Natural Language Processing (EMNLP)*.
- Soujanya Poria, Alexander Gelbukh, Amir Hussain, Newton Howard, Dipankar Das, and Sivaji Bandyopadhyay. 2013. Enhanced sentinet with affective labels for concept-based opinion mining. *IEEE Intelligent Systems* 28(2):2–9.
- Ines Rehbein, Josef Ruppenhofer, and Caroline Sporleder. 2009. Assessing the benefits of partial automatic pre-labeling for frame-semantic annotation. In *Proc. of the Third Linguistic Annotation Workshop*. Association for Computational Linguistics, Suntec, Singapore, pages 19–26.
- Sascha Rothe, Sebastian Ebert, and Hinrich Schütze. 2016. Ultradense word embeddings by orthogonal transformation. *CoRR* abs/1602.07572.
- Carlo Strapparava and Rada Mihalcea. 2007. Semeval-2007 task 14: Affective text. In *Proc. of the 4th International Workshop on Semantic Evaluations*. pages 70–74.