In Search for Lost Emotions: Deep Learning for Opinion Taxonomy Induction

Elena Melnikova[1], Emmanuelle Dusserre[2] and Muntsa Padró[2]

[1] Innoradiant, Meylan 38240, France
[2] Eloquant, Gières 38610, France
elena.melnikova@innoradiant.com
emmanuelle.dusserre@eloquant.com
muntsa.padro@eloquant.com

Abstract. In this article, we present an approach for using word2vec to automatically enrich the opinions' taxonomy used by a sentiment analysis system. More specifically, we worked on emotion lexicon for the field of customer relationship management. The proposed method consists of searching for the nearest distributional neighbors of each source word, and add them to the lexicon of emotions. The hypothesis is that the contextual neighbors of the emotions will also carry an emotional coloring. The results of this experiment show that the neighborhood lexicon is not sufficiently representative. Nevertheless, most of the collected items seem to express another type of opinion, namely judgments. This unexpected result allows us to broaden our taxonomy of opinions with a new informational field, richer and more expressive.

Keywords: emotions, judgments, opinions, taxonomy, word2vec, deep learning, sentiment analysis.

1 Introduction

The automatic analysis of customer opinions is becoming one of the most pervasive concerns of companies studying customer reviews. The sentiments or opinions expressed in these reviews are important indicators for the company's decision-making strategy. Thus, sentiment analysis (SA) systems need to be reliable and constantly updated.

Many SA systems are often based on social networks, tweets or SMS corpora [6, 1, 3]. The analysis of opinions essentially focuses on polarity detection in customers feedbacks (positive, negative or neutral), [4]. Our SA system for French [7] performs the extraction of different kind of fine-grained opinions, including emotions, to extract more detailed information than just positive vs negative polarity. This fine-grained detection is performed using a taxonomy associating words or expressions to the classes to be detected [11]. Though, building a complete taxonomy can be very time consuming, and the relevant terms might depend on the domain.

The present work is motivated by the wish to semi-automatically enrich the taxonomy, in order to achieve a greater accuracy and easily adapt the system to different sub-domains. To do so, we propose the use of word2vec [10] to add entries to an existing taxonomy. This method has been widely and successfully used in semantic analysis and other NLP tasks [2, 9]. We assume that the matrix constructed by traversing our domain specific corpus (Customer Relationship Management or CRM) would locate close to each other emotional words belonging to the same class (from the distributional point of view). By adopting this method based on deep learning, we intend to check if it is relevant for the enrichment of our emotions taxonomy.

The rest of this paper is structured as follows. In Section 2, we describe the word2vec method and the procedure to enrich the emotion taxonomy. Section 3 reports experiments and result discussion. Finally, Section 4 concludes our work.

2 The word2vec method

Word2vec is a statistical language model based on neural networks developed by a team of researchers under the direction of T. Mikolov [10] at Google[1]. This method is used to produce word embeddings: words in a corpus are represented as vectors in a multidimensional space, their position in this space corresponding to their semantic representation [8, 10].

The word2vec technique is based on a distributional hypothesis: words that appear in the same context share semantic values, thus, words that are close in the space are semantically close. Our goal is to use this method to find sets of closest words (or related words) and assign them to the same semantic class to enrich other kind of taxonomies, as it was shown in [2, 5].

We followed a procedure which contains two major steps. The first one is to create a word2vec model from a given corpus. The model creation includes the processing of the corpus (tokenization, lemmatization, etc.) and the induction of the model from it. The second step is to use the model to calculate the distance between a selected word (*seed* word) and the other words in the corpus. In our work, the words already present in the taxonomy (that already have an assigned semantic tag) serve as seed words. The objective is to assign to the closest words of the seed words the same semantic class as them. The distance between two words is calculated with the cosine of the angle between the vectors that represent them. The more this cosine is close to 1, the closest the neighbor is to the source word.

3 Experiments and results

We first focused on the improvement of the emotion detection module, by reviewing the emotions taxonomy used by our system. This taxonomy contains lexical items from different linguistic genres[2] (literature, psychology, familiar) with 41 classes and more than 1100 words. This seemed too large and it was not adapted to our CRM domain. Thus, the taxonomy was considerably reduced to constitute a specific sub-taxonomy (10 classes, 360 words). Among these classes, we find: ANGER, SADNESS, DISSATISFACTION, LIKING, SATISFACTION, DISLIKE, DOUBT, TRUST, CALMNESS. The lexicon for this reduced classification seemed quantitatively "poor" and not adequate to the CRM domain. Thereby, we found it necessary to increase the number of words

[1] https://code.google.com/archive/p/word2vec/

[2] WordNet based taxonomy : https://wordnet.princeton.edu/wordnet/frequently-asked-questions/database/ (Princeton University 2018)

for certain classes that contained less than 10 items (SATISFACTION, TRUST, DISTRUST, DOUBT, DISLIKE, CALMNESS).

The corpus used for the model has 15 million words and it is very specific to the CRM domain. Despite word2vec models are expected to work better with bigger corpus, [5] showed that for specific domains it is preferable to have a domain specific corpus than huge amount of data.

Table 1 shows a sample of results obtained when applying the word2vec method for a selection of words of emotion classes poorly endowed. The headline shows the seed words with their emotion classes and the following lines, the neighbors proposed with their cosine.

Table 1. Extraction of the nearest neighbors

Confiance (trust) class: TRUST		***Satisfait*** (satisfied) class: SATISFACTION		***Doute*** (doubt) class: DOUBT	
Neighbours	**Cos.**	**Neighbours**	**Cos.**	**Neighbours**	**Cos.**
perdre (to lose)	0.46	*résilier* (to cancel)	0.31	*red* (red)	0.39
préférer (to prefer)	0.41	*resolu* (resolved)	0.31	*prévenir* (to prevent)	0.37
abus (breach)	0.39	*accueil* (reception)	0.32	*accord* (agreement)	0.38
quitter (to leave)	0.30	*satisfaire* (to satisfy)	0.32	*considération* (consideration)	0.34

This extract is obtained by using a threshold of 0.3, meaning that only words with a cosine bigger than 0.3 are suggested as candidates to be added to the taxonomy. The threshold is set heuristically, looking for a compromise where we retrieve enough candidates without too much noise. In the obtained results we noticed different types of distributional neighborhoods:

1. collocations (*perdre confiance* (lose confidence), *abus de confiance* (breach of trust), *gagner en (la) confiance (de qqn)* (gain someone's trust), *satisfait de l'accueil* (satisfied with the reception));
2. synonyms, antonyms and derived forms (*satisfaire* (to satisfy) for *satisfait* (satisfied));
3. neighbors of a different semantic tag (*code* (code), *œuvre* (work), *cordialement* (cordially), *fixer* (to fix), *supprimer* (to delete), etc.).

The lexica that we expected to find should come from the second type of neighborhood, according to our goal and hypothesis. The obtained results show that these lexica are rather infrequent and, most often, they contain antonyms. It means that we cannot increase the classes of our taxonomy by using the closest neighbors. Thus, our hypothesis of enriching the emotions taxonomy, and especially poorly endowed emotions classes by using word2vec is not confirmed. Nevertheless, the idea of resorting to a new method which provides a rich lexical and statistical information, seems to us very attractive and less exploited.

3.1 The unexpected result

To better understand the results, we study more closely the lexica coming from the most numerous type of neighborhood, the neighbors with different semantic tag.

Some of those neighbors can be considered just noise, but we have distinguished in this lexical layer a category of words that could characterize non-emotional opinions, the *judgments.* For example, the words like *beau* (beautiful) and *accord* (agreement) are positive polarity judgments. The words *proche* (close adj.), *rapide* (fast) are positive or negative depending on the domain[3]. This type of polar judgment lexicon is as important as the emotion lexicon for the detection and analysis of customers opinions.

At the sight of these results, we extended the experiment by using word2vec with the whole taxonomy of emotions as seed words. This allowed us to identify more words that are likely to express judgments (see Table 2).

Table 2. Extraction of the nearest neighbors to search for the lexicon of judgments.

Déplaisant (unpleasant) **Class: DISLIKE**		***Rassurant*** (reassuring) **Class: TRUST**		***Satisfait*** (satisfied) **Class: SATISFACION**	
Neighbours	**Cos.**	**Neighbours**	**Cos.**	**Neighbours**	**Cos.**
Serviabilité (helpfulness)	0.28	***Comprehensible*** (understandable)	0.29	*Traitement* (proscessing)	0.23
Relation (relationship)	0.26	*Approprier* (to appropriate)	0.26	***Courtois*** (courteous)	0.23
Impeccable (faultless)	0.21	*Plateforme* (platform)	0.26	***Performant*** (performing)	0.20

Table 2 shows in bold the neighbors that are judgments and that can be added to our lexica[4]. In fact, our SA system already contains a little judgment lexicon (13 classes with 197 items). The new results can be gathered in a class APPRECIATION completing this lexicon. Note that the selection of this class for the judgment words is done manually, and not assigning the semantic class of the seed word, as it was our hypothesis. Even though it demands more human intervention, the possibility of enriching the judgment taxonomy seems to us a promising outcome of this work, since by studying the output of word2vec we can find new relevant classes. Word2vec finds more lexica related to judgments than emotions since CRM corpora contain much more of these lexica. This is also why judgments constitute key elements to be added to our SA system.

3.2 The extension of the judgments taxonomy

The judgments lexica extracted from CRM corpus better characterizes the CRM-specific classes (see Fig. 1). This idea is corroborated by a recent research study [5]. The augmented judgment taxonomy contains 18 classes and 261 items compared to 13 classes and 197 words from the initial taxonomy of judgments. Table 3 shows an extract from this ranking that should be expanded and completed.

[3] Their polarity reveals in context (*le personnel est proche du client* (the staff is close to the client*)* [positive judgment in commercial context vs negative judgment in familiar context]).

[4] In this second experiment we use 0.2 as threshold in order to obtain more candidates to be added to the judgment taxonomy

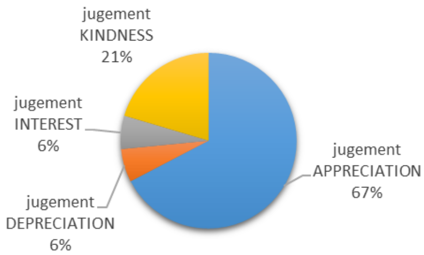

Fig. 1. The distribution of the lexicon of judgments extracted from CRM corpus.

Table 3. An extract from the new taxonomy of judgments.

Opinion's type	**Class**	**Words**
judgment	APPRECIATION	*Positive* (positive) *Apprécier* (to appreciate) *Méritoire* (meritorious)
	DEPRECIATION	*Déprécié* (depreciated) *Fichu* (damn) *Anormal* (unnatural)

4 Conclusion and Further Work

In this work, we have tested the applicability of word2vec to enrich an existent emotion taxonomy by finding semantically close words. The obtained results are not very satisfactory, since very few new emotional words are extracted. Nevertheless, the method allowed us to extract judgment words which are also very important for the SA system and much more frequent in our domain. Thus, we can use this new taxonomy as a resource in our system and we conclude that word2vec is a useful method to enrich existing taxonomies and even to discover new classes. But to guarantee the quality of final resources human intervention is needed. Table 4 summarizes the most important positive and negative points we spotted with our experiments with word2vec.

Table 4. Advantages and disadvantages of word2vec

Criteria	**Advantages**	**Disadvantages**
Speed	Tailor made extraction (depending on the threshold, PoS filter, etc.)	The learning time can take several tens of minutes depending on the memory size of the machine and corpus size.
Efficiency	The matrix can be trained with several types of corpus	Ambiguous lexica is often present, which requires to check its meaning in the context to associate it to a class or to remove it from the taxonomy
Reliability	The use of a specific domain corpus improves the results	We cannot predict the type of extracted lexica (expected affects, harvested judgments)

In the future, we plan to perform an extrinsic evaluation of the developed taxonomies. The convenience of using word2vec to enrich the taxonomies seems clear to us, when considering as an alternative the manual development of the taxonomies. Nevertheless, a final evaluation of our SA system before and after enriching the taxonomy needs to be done. For that, we are currently working on a CRM gold-standard. Also, we plan to perform the same experiments with word2vec trained on another CRM corpus to adapt the taxonomy of its sub-domain.

Furthermore, the taxonomy enriched with word2vec can also serve as input for a new run of the taxonomy enrichment system. Thus, we could perform a bootstrap to iteratively enrich the taxonomy of emotions and judgments.

References

1. Abdaoui, A., Nzali, M.D.T., Azé, J., Bringay, S., Lavergne, Ch., et al. ADVANSE: Sentiment, Opinion and Emotion Analysis in French Tweets. DEFT: Défi Fouille de Texte, Jun 2015, Caen, France. Actes de la 11e Défi Fouille de Texte (2015) <hal-01222629>
2. Baroni, M., Dinu, G. & Kruszewski, G. Don't count, predict! A systematic comparison of context-counting vs. contextpredicting semantic vectors. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 238–247, Baltimore, Maryland, June. Association for Computational Linguistics (2014).
3. Dini, L., Bittar, A., Robin, C., Segond, F., Montaner, M., SOMA: The Smart Social CRM. Handling Semantic Variability of Emotion Analysis with Hybrid Technologies. In Sentiment Analysis in Social Network Elsevier, chapter 13 (2016).
4. Dridi, A., Reforgiato Recupero, D. Leveraging semantics for sentiment polarity detection in social media. International Journal of Machine Learning and Cybernetics (2017).
5. Dusserre, E. & Padró, M. Bigger does not mean better ! We prefer specificity. In 12th International Conference on Computational Semantics (IWCS). 19-22 September 2017 Montpellier (France) (2017).
6. Hamon, T., Fraisse, A., Paroubek, P., Zweigenbaum, P., Grouin, C.. Analyse des émotions, sentiments et opinions exprimés dans les tweets: présentation et résultats de l'édition 2015 du défi fouille de texte (DEFT). In 22ème Traitement Automatique des Langues Naturelles (2015).
7. Maurel, S., Curtoni, P., & Dini, L. A hybrid method for sentiment analysis. In INFORSID. Présenté à Défi Fouille de Texte 2007 (DEFT'07) (2008).
8. Levy, O., Goldberg, Y. & Dagan, I: Improving Distributional Similarity with Lessons Learned from Word Embeddings. Transactions of the Association for Computational Linguistics (2015).
9. Maître, J., Menard, M., Chiron, G., Bouju, A . Utilisation conjointe LDA et Word2Vec dans un contexte d'investigation numérique. Extraction et Gestion des Connaissances 2017, Jan 2017, Grenoble, France (2017).
10. Mikolov T., Sutskever I., Chen K., Corrado G., Dean G. Distributed Representations of Words and Phrases and their Compositionality. NIPS'13 Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada (2013).
11. Whitelaw, C., Garg, N., Argamon, S. Using Appraisal Taxonomies for Sentiment Analysis. Conference: The Second Midwest Computational Linguistic Colloquium, MCLC (2005).