RECOGNIZING EMOTIONS IN SHORT TEXTS

Ovidiu Şerban[1,2], Alexandre Pauchet[1] and Horia F. Pop[2]
[1]*LITIS, INSA de Rouen, Avenue de l'Université - BP 8, 76801 Saint-Étienne-du-Rouvray, France*
[2]*Fac. of Mathematics & Computer Science, Babeş-Bolyai University, 1, M. Kogălniceanu St., 400084 Cluj-N., Romania*

Keywords: Natural language processing, Machine learning, Affective computing, Text mining, Emotion detection.

Abstract: Affective Computing is one of the fields used by computer scientists to transfer the knowledge from psychology to the Human-Machine Interaction research field, while offering a better understanding on Human to Human Interaction. Since the classification problem is not typical, the difficulty is increased by the fuzziness of the data sets. Our paper proposes a method that aims at a better recognition rate of human emotions. Our model is based on the Self-Organizing Maps algorithm and it can be applied on short texts with a high degree of affective content. It is designed to be integrated into an Embodied Conversational Agent.

1 INTRODUCTION

Emotion detection has been widely approached by different anthropologists and psychologists (Calvo and D'Mello, 2010), starting with Charles Darwin (Darwin, 1872) who considered that emotions are universal (i.e. identical for humans and animals). Later, W. James (James, 1884) and P. Ekman (Ekman et al., 1998), extended Darwin's theory, but they retained the concept of affective universality.

In computer science, emotion detection is proposed as a solution for the challenge of human-computer interactions and it has been tackled by projects (e.g. SEMAINE (Schroder, 2010)), which aims at creating an Embodied Conversational Agent able to detect simple emotions and sustain interaction with the user through affective features in the agent's language and behaviour.

While detection of emotional states tends to be approached by classical Machine Learning techniques (Calvo and D'Mello, 2010; Picard, 2000), the problem of affective behaviour simulation is tackled by groups that developed Affective Embodied Conversational Agents (e.g. Greta (Pelachaud, 2009)). Both detection and simulation can be studied through the perspective of Affective Computing.

Objective. Emotion detection is increasingly used in Embodied Conversational Agents to create an adapted reply channel to the user's affective state. In this context, we propose a method to detect emotions in short texts (i.e. in texts whose size is similar to dialogue utterances). Our goal is to design a model to detect the dominant affective state produced by short texts onto a reader and to classify them into six clusters, corresponding to Ekman's psychological theory.

In the current paper, the corpus consists of newspaper headlines, from SemEval 2007, task 14 (Strapparava and Mihalcea, 2008). The corpus was chosen because of the appropriate size of its elements and their high emotional content. Since the methods presented in the paper, related to the corpus do not offer a good accuracy, we introduce a new classification mechanism based on the Self Organizing Maps. Also, our approach can be easily transposed to other contexts such as chat logs, forums or oral transcripts.

This paper is organized as follows: the next paragraph describes the related work, in Section 2 we make a short presentation of the corpus we are working on, followed by more details of our method. Afterwards, in Section 3 we describe some results we obtained and finally we conclude and present the future work in Section 4.

Related Work. Several experiments were carried out from a corpus evaluation perspective, like the one presented in (Calvo and D'Mello, 2010). All the approaches can be classified into two main categories: 1) approaches that use ontologies or word databases (e.g. WordNet synsets) to distinguish between classes of emotions and 2) specialised approaches.

As a synset database example, we will mention WordNet Affect (Strapparava and Valitutti, 2004), an extension of the WordNet data set. WordNet Affect is a 6 class annotation (i.e. Ekman's basic annotation scheme) made on a synset level. Also, SentiWordNet

Şerban O., Pauchet A. and F. Pop H..
RECOGNIZING EMOTIONS IN SHORT TEXTS.
DOI: 10.5220/0003718004770480
In *Proceedings of the 4th International Conference on Agents and Artificial Intelligence* (ICAART-2012), pages 477-480
ISBN: 978-989-8425-95-9

(Baccianella et al., 2010) is the result of automatic annotation of all WordNet synsets according to their degrees of positivity, negativity, and neutrality.

Starting from WordNet Affect, (Valitutti et al., 2005) proposed a simple word presence method to detect emotions. (Ma et al., 2005) designed an emotion extractor from chat logs, based on the same simple word presence. SemEval 2007, task 14 (Strapparava and Mihalcea, 2008) presented a corpus and some methods to evaluate it, some based on Latent Semantic Analyser (LSA) and presence of an emotional word (e.g. WordNet Affect item).

Methods more related to signal processing were proposed by (Alm et al., 2005), (Danisman and Alpkocak, 2008), or (D'Mello et al., 2006) which introduce different solutions for feature extraction and selection and various classifiers. (Alm et al., 2005) used a corpus of child stories and a Winnow Linear method to classify the data into 7 categories. Using the ISEAR (Wallbott et al., 1988) dataset, a popular collection of psychological data from around 1990, (Danisman and Alpkocak, 2008) used different classifiers like Vector Space Model (VSM), Support Vector Machine (SVM) or a Naive-Bayes (NB) method to distinguish between 5 categories of emotions.

2 EMOTION CLASSIFICATION

Emotional Corpus. The chosen corpus for our experiment is from SemEval 2007, task 14 (Strapparava and Mihalcea, 2008), proposed at the conference with the same name. The data set contains headlines (newspaper titles) from major websites, such as New York Times, CNN, BBC or Google News.

The corpus was manually annotated by 6 different persons. They were instructed to annotate the headlines with emotions according to the presence of affective words or group of words with emotional content. The annotation scheme used for this corpus is the basic six emotions set, presented by Ekman: Anger, Disgust, Fear, Joy (Happiness), Sadness, Surprise. In situations were the emotion was uncertain, they were instructed to follow their first feeling. The data is annotated with a 0 to 100 scale for each emotion.

The authors of the corpus proposed a double evaluation, on a fine-grained scale and on a coarse-grained scale. For the fine-grained scale, for values from 0 to 100, the system results are correlated using the Pearson coefficients described by the inter-annotator agreement. The second proposition was a coarse-grained encoding, where every value from the 0 to 100 interval is mapped to either 0 or 1 (0 =[0,50) , 1=[50,100]). Considering the coarse-grained evaluation, a simple overlap was performed.

Classification Model. The classifier we have chosen is a commonly used unsupervised method, the Self-Organizing Maps (SOM) (Kohonen, 1990). This method is a particular type of neural network used for mapping large dimensional spaces into small dimensional ones. The SOM has been chosen because: 1) it usually offers good results with fuzzy data, 2) the training process is easier than other Neural Networks and 3) the classification speed is sufficiently high.

Preprocessing Step. During the preprocessing step, we applied on each headline a collection of filters, in order to remove any useless information, such as special characters and punctuation, camel-case separators and stop word filtering[1].

This method offers a good balance between speed and accuracy of the results, compared to other methods like Part of Speech Tagging (POS), which provides comparable results, but tends to be slower.

Feature Extraction. We have chosen LSA, applied with three different strategies. Hence, all the occurrences of key terms are counted and introduced to a matrix (a row for each keyword, a column by headline). The term set (keywords) is chosen according to three different strategies.

The first LSA strategy we implemented concerns the algorithm applied onto the words of the WordNet Affect database (Strapparava and Valitutti, 2004). This method is called pseudo-LSA or meta-LSA by C. Strapparava and R. Mihalcea (Strapparava and Mihalcea, 2008). The meta-LSA algorithm differs from the classic implementation by using clusters of words instead of single words. This strategy did not provide the expected results: the recall decreased since all of the presented words were carrying an emotional value and the non-emotional words were not represented. Our version confirms the results obtained by Mihalcea and Strapparava.

The second strategy use the classic LSA applied onto the words of the training set. While the genericness of this approach is not assured by the support word collection, this method offers a good starting point for similar training corpus and testing corpus.

Our third proposition was to use the top 10 000 most frequent English words, extracted from approximately 1 000 000 documents existing in the Project Gutenberg[2]. The features used are the document sim-

[1] We considered as stop words all prepositions, articles and other short words that do not carry any semantic value (e.g. http://www.textfixer.com/resources/common-english-words.txt)

[2] Project Gutenberg is a large collection of e-books, processed and reviewed by the project's community.

ilarities obtained after applying the LSA algorithm.

Feature Selection. After the feature extraction, the feature selection is performed by using a k-LSA[3] instead of the classical version of the algorithm, because this algorithm reduces the feature space by removing the ones which would not aid the classification.

SOM. Many of the proposed implementations of the Self-Organizing Maps use the feature model or a linear combination of the features for classification. Our implementation is very close to the classical ones, but the feature space and classes were split into two distinct concepts and the classes are not used actively in the self-organizing algorithm; **data** and **label** vectors are separated in the Self-Organized Nodes and the learning process is done similarly for both of the vectors, with the same parameters.

A 40x40 grid size was used for the SOM configuration. The feature vectors were the document similarity vectors obtained from the feature extraction step, i.e. the columns of the V^T matrix computed in the SVD decomposition from the LSA algorithm. As for the labels, we used the intensities available in the corpus as an independent vectorial space.

Classification. For the classification part, we used the same measure as during the training phase, which computes a distance from a proposed individual to all the elements in the SOM grid. The Best Matching Unit is selected, i.e. the element of the grid which is closest to the desired individual. In our experiments, the Euclidean distance was used both in the SOM algorithm and for evaluation.

3 RESULTS

During the SemEval 2007 task, the coarse-grained evaluation did not provide the expected results. Therefore, we started with two experiments in order to discover any kind of class dominance. Firstly, only the emotional values were taken into consideration, but this approach failed to extract any dominant class. Secondly, the neutral class (No Emotion) was added, leading to an important result, as shown in Table 1. The neutral class is observed with a strong dominance over the other classes, i.e. 64 % dominant value. The conclusion of this experiment is that neither of the classifiers presented at the SemEval 2007 conference managed to break the dominance of the neutral class, and the classifier we proposed discovers the neutral class better than the others.

Table 1: Dominant class for coarse-grained representation.

	Nb. of instances	
No emotion	642	64.85%
Anger	14	1.41%
Disgust	6	0.61%
Fear	65	6.57%
Joy	110	11.11%
Sadness	81	8.18%
Surprise	38	3.84%
Combined	34	3.43%

The second experiment concerns the whole corpus, with a coarse-grained representation. All the results are presented in Table 3. The LSA training column represents the LSA decomposition method applied on the words extracted from the training corpus, while the LSA Gutenberg column presents the results of the k-LSA method applied on the 10 000 words extracted from the Gutenberg corpus. Among our models, we present the most significant scores obtained by the systems participating in the SemEval 2007, task 14 competition (Strapparava and Mihalcea, 2008). Also, we present the overall (Table 2).

Table 2: Overall results.

	Precision	Recall	F1
LSA training	20.50	19.57	20.02
LSA Gutenberg	24.22	23.31	**23.76**
LSA All emotion	9.77	**90.22**	17.63
UA	17.94	11.26	13.84
UPAR7	**27.60**	5.68	9.42

The results are not surprising, since LSA All emotions offers a good coverage over the emotional words, but its synonym expansion algorithm introduces noise in the method, and therefore offers a very poor precision. UPAR7 leads in some cases to a good precision, due to its analytical nature, but it lacks in recall. Our system offers a good compromise between precision and recall, as the F1 measure shows.

4 CONCLUSIONS

We present a method for recognizing emotions in short texts, designed to be integrated into an Embodied Conversational Agent. In other words, the length of the analysed texts corresponds to the length of utterances during a dialogue. Our model, based on LSA and a SOM algorithm, benefits from the power of unsupervised neural networks, which obtain better results on fuzzy data and which propose an easy-to-perform training step.

All the documents are freely available at the website: http://www.gutenberg.org/wiki/Main_Page

[3]The k-LSA version eliminates the null values from the Σ diagonal matrix and k is the reduction index

Table 3: The systems presented in the SemEval competition.

	LSA training			LSA Gutenberg			LSA All emotional			UA			UPAR7		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
A.	10.00	11.86	10.85	**18.52**	15.38	**16.80**	6.20	**88.33**	11.59	12.74	21.60	16.03	16.67	1.66	3.02
D.	3.33	4.17	3.70	**8.33**	7.69	**8.00**	1.98	**94.12**	3.88	0.00	0.00	-	0.00	0.00	-
F.	19.01	17.76	18.36	28.39	27.67	**28.03**	12.55	**86.44**	21.92	16.23	26.27	20.06	**33.33**	2.54	4.72
J.	36.75	36.75	36.75	40.49	64.62	**49.79**	18.60	**90.00**	30.83	40.00	2.22	4.21	**54.54**	6.66	11.87
Sa.	24.14	40.00	30.11	27.08	19.60	22.74	11.69	**87.16**	20.62	25.00	0.91	1.76	**48.97**	22.02	**30.38**
Su.	**29.73**	6.92	11.23	22.50	4.95	8.11	7.62	**95.31**	14.11	13.70	16.56	**14.99**	12.12	1.25	2.27

Anger=**A**, Disgust=**D**, Fear=**F**, Joy=**J**, Sadness=**Sa.**, Surprise=**Su.**

The linguistic part of our model, with most frequent used words in English, offers a good score on global F1 and a good global precision, better than most models tested on this corpus. Even if this linguistic model should be limited in certain situations, it provides a good image over the English language. Moreover, it can be built faster than most models.

As a proposition, we intend to improve our linguistic model with a different support words in order to represent better emotional contents. In that way, we plan to build an alternative dictionary able to discover new emotional words in relation with their context, and could improve the current classification method.

Besides, in order to increase the genericness of the system, we intend to extend the training base with several existing corpuses collected and validated during a real-time and real-life experiment. One of the ways to obtain such an integration is through an Affective Embodied Conversational Agent as a tutoring partner for a generic task.

REFERENCES

Alm, C., Roth, D., and Sproat, R. (2005). Emotions from text: machine learning for text-based emotion prediction. In *Proc. of the conf. on Human Lang. Technology and Empirical Methods in NLP*, pages 579–586. Assoc. for Comp. Linguistics.

Baccianella, S., Esuli, A., and Sebastiani, F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Seventh conf. on Int. Lang. Res. and Eval., Malta. Retrieved May*, volume 25, page 2010.

Calvo, R. and D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, pages 18–37.

Danisman, T. and Alpkocak, A. (2008). Feeler: Emotion classification of text using vector space model. In *AISB 2008 Convention Communication, Interaction and Social Intelligence*, volume 1, page 53.

Darwin, C. (1872). The expression of emotions in animals and man. *Nueva York: Appleton. Traducción*.

D'Mello, S., Craig, S., Sullins, J., and Graesser, A. (2006). Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. *Int. Journal of AI in Education*, 16(1):3–28.

Ekman, P., Friesen, W., JENKINS, J., OATLEY, K., and STEIN, N. (1998). Constants across cultures in the face and emotion. *Human emotions*, pages 63–72.

James, W. (1884). What is an Emotion? *Mind*, 9(34):188–205.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480.

Ma, C., Prendinger, H., and Ishizuka, M. (2005). A chat system based on emotion estimation from text and embodied conversational messengers. *Entertainment Computing-ICEC 2005*, pages 535–538.

Pelachaud, C. (2009). Modelling multimodal expression of emotion in a virtual agent. *Philosophical Trans. of the Royal Society B: Biological Sciences*, 364(1535):3539.

Picard, R. (2000). *Affective computing*. The MIT press.

Schroder, M. (2010). The semaine api: towards a standards-based framework for building emotion-oriented systems. *Advances in HCI*, 2010:2–2.

Strapparava, C. and Mihalcea, R. (2008). Learning to identify emotions in text. In *Proc. of the 2008 ACM symposium on Applied computing*, pages 1556–1560. ACM.

Strapparava, C. and Valitutti, A. (2004). WordNet-Affect: an affective extension of WordNet. In *Proceedings of LREC*, volume 4, pages 1083–1086. Citeseer.

Valitutti, A., Strapparava, C., and Stock, O. (2005). Lexical resources and semantic similarity for affective evaluative expressions generation. *Affective Computing and Intelligent Interaction*, pages 474–481.

Wallbott, H., Scherer, K., et al. (1988). Emotion and economic developmentData and speculations concerning the relationship between economic factors and emotional experience. *European journal of social psychology*, 18(3):267–273.