Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence

AffectiveSpace 2: Enabling Affective Intuition for Concept-Level Sentiment Analysis

Erik Cambria
School of Computer Engineering
NTU, Singapore
cambria@ntu.edu.sg

Jie Fu
School of Computing
NUS, Singapore
jie.fu@comp.nus.edu.sg

Federica Bisio
DITEN
University of Genoa, Italy
federica.bisio@edu.unige.it

Soujanya Poria
School of Computer Engineering
NTU, Singapore
sporia@ntu.edu.sg

Abstract

Predicting the affective valence of unknown multi-word expressions is key for concept-level sentiment analysis. AffectiveSpace 2 is a vector space model, built by means of random projection, that allows for reasoning by analogy on natural language concepts. By reducing the dimensionality of affective common-sense knowledge, the model allows semantic features associated with concepts to be generalized and, hence, allows concepts to be intuitively clustered according to their semantic and affective relatedness. Such an affective intuition (so called because it does not rely on explicit features, but rather on implicit analogies) enables the inference of emotions and polarity conveyed by multi-word expressions, thus achieving efficient concept-level sentiment analysis.

Introduction

Concept-level sentiment analysis focuses on a semantic analysis of text through the use of web ontologies or semantic networks, which allow the aggregation of conceptual and affective information associated with natural language opinions. By relying on large semantic knowledge bases, such approaches step away from blind use of keywords and word co-occurrence count, but rather rely on the implicit features associated with natural language concepts.

Unlike purely syntactical techniques, concept-based approaches are able to detect also sentiments that are expressed in a subtle manner, e.g., through the analysis of concepts that do not explicitly convey any emotion, but which are implicitly linked to other concepts that do so. The bag-of-concepts model can represent semantics associated with natural language much better than bags of words. In the bag-of-words model, in fact, a concept such as `cloud computing` would be split into two separate words, disrupting the semantics of the input sentence (in which, for example, the word `cloud` could wrongly activate concepts related to `weather`).

The analysis at concept-level allows for the inference of semantic and affective information associated with natural language opinions and, hence, enables a comparative fine-grained feature-based sentiment analysis. Rather than gathering isolated opinions about a whole item (e.g., iPhone6), users are generally more interested in comparing different products according to their specific features (e.g., iPhone6's vs GalaxyS6's touchscreen), or even sub-features (e.g., fragility of iPhone6's vs GalaxyS6's touchscreen).

In this context, common-sense knowledge is key for properly deconstructing natural language text into sentiments – for example, to appraise the concept `small room` as negative for a hotel review and `small queue` as positive in a patient opinion, or the concept `go read the book` as positive for a book review but negative for a movie review. The inference of emotions and polarity from natural language concepts, however, is a formidable task as it requires advanced reasoning capabilities such as common-sense, analogical, and affective reasoning.

In this work, we present AffectiveSpace 2, a novel vector space model for concept-level sentiment analysis that allows for reasoning by analogy on natural language concepts, even when these are represented by highly dimensional semantic features. In this sense, AffectiveSpace 2 can be seen as a powerful tool to address the emerging issue of "Big Dimensionality" (Zhai, Ong, and Tsang 2014) in the context of natural language processing (NLP) and sentiment analysis. The proposed model, however, should not be considered solely as a NLP tool, but rather as a framework for analogical reasoning that can be embedded in potentially any cognitive system dealing with real-world semantics, e.g., concepts associated with images (Cambria and Hussain 2012a), audio (Principi et al. 2015), handwriting (Wang et al. 2013), and multimodal data (Poria et al. 2015).

The rest of the paper is organized as follows: the first section presents related work in the field of concept-level sentiment analysis; the following two sections describe in detail how AffectiveSpace 2 is built and clustered, respectively; an evaluation section proposes experimental results for an opinion mining task; finally, the last section provides some concluding remarks.

Related Work

Concept-level sentiment analysis is a NLP task that has recently raised growing interest both within the scientific community, leading to many exciting open challenges, as well as in the business world, due to the remarkable benefits to be had from marketing and financial market prediction.

The potential applications of concept-level sentiment analysis, in fact, are countless and span interdisciplinary areas such as stock market prediction, political forecasting, social network analysis, social stream mining, and human-robot interaction.

For example, Li et al. (Li et al. 2014) implemented a generic stock price prediction framework and plugged in six different models with different analyzing approaches. They used Harvard psychological dictionary and Loughran-McDonald financial sentiment dictionary to construct a sentiment space. Textual news articles were then quantitatively measured and projected onto such a sentiment space. The models' prediction accuracy was evaluated on five years historical Hong Kong Stock Exchange prices and news articles and their performance was compared empirically at different market classification levels.

Rill et al. (Rill et al. 2014) proposed a system designed to detect emerging political topics in Twitter sooner than other standard information channels. For the analysis, authors collected about 4 million tweets before and during the parliamentary election 2013 in Germany, from April until September 2013. It was found that new topics appearing in Twitter can be detected right after their occurrence. Moreover, authors compared their results to Google Trends, observing that the topics emerged earlier in Twitter than in Google Trends.

Jung and Segev (Jung and Segev 2014) analyzed how communities change over time in the citation network graph without additional external information and based on node and link prediction and community detection. The identified communities were classified using key term labeling. Experiments showed that the proposed methods can identify the changes in citation communities multiple years in the future with performance differing according to the analyzed time span.

Montejo-Raez et al. (Montejo-Raez et al. 2014) introduced an approach for sentiment analysis in social media environments. Similar to explicit semantic analysis, microblog posts were indexed by a predefined collection of documents. In the proposed approach, performed by means of latent semantic analysis, these documents were built up from common emotional expressions in social streams.

Bell et al. (Bell et al. 2014) proposed a novel approach to social data analysis, exploring the use of microblogging to manage interaction between humans and robots, and evaluating an architecture that extends the use of social networks to connect humans and devices. The approach used NLP techniques to extract features of interest from textual data retrieved from a microblogging platform in real-time and, hence, to generate appropriate executable code for the robot. The simple rule-based solution exploited some of the 'natural' constraints imposed by microblogging platforms to manage the potential complexity of the interactions and to create bi-directional communication.

Building AffectiveSpace 2

The best way to solve a problem is to already know a solution for it. But, if we have to face a problem we have never met before, we need to use our intuition. Intuition can be explained as the process of making analogies between the current problem and the ones solved in the past to find a suitable solution. This kind of thinking is maybe the essence of human intelligence since in everyday life no two situations are ever the same and we have to continuously perform analogical reasoning for problem solving and decision making.

The human mind constructs intelligible meanings by continuously compressing over vital relations (Fauconnier and Turner 2003). The compression principles aim to transform diffuse and distended conceptual structures to more focused versions so as to become more congenial for human understanding. In order to emulate such a process, principal component analysis (PCA) was previously applied on the matrix representation of AffectNet (Cambria and Hussain 2012b), a semantic network in which common-sense concepts were linked to semantic and affective features (Table 1). The result was AffectiveSpace.

PCA is most widely used as a data-aware method of dimensionality reduction (Jolliffe 2005). PCA is closely related to the low-rank approximation method, singular value decomposition (SVD), in the sense that PCA works on a transformed version of the data matrix (Menon and Elkan 2011). SVD seeks to decompose the AffectNet matrix $A \in \mathbb{R}^{n\times d}$ into three components,

$$A = USV^T, \tag{1}$$

where U and V are unitary matrices, and S is an rectangular diagonal matrix with nonnegative real numbers on the diagonal.

SVD has been proved to be optimal in preserving any unitarily invariant norm[1] $\| \cdot \|_M$ (Menon and Elkan 2011):

$$\| A - A_k \|_M = \min_{\mathrm{rank(B)=k}} \| \mathrm{A} - \mathrm{B} \|_{\mathrm{M}}, \tag{2}$$

where A_k, i.e., AffectiveSpace, is formed by only containing the top k singular values in S. Hence, in AffectiveSpace, common-sense concepts and emotions are represented by vectors of k coordinates. These coordinates can be seen as describing concepts in terms of 'eigenmoods' that form the axes of AffectiveSpace, i.e., the basis e_0,...,e_{k-1} of the vector space. For example, the most significant eigenmood, e_0, represents concepts with positive affective valence. That

[1] A norm $\|\cdot\|_M$ is unitarily invariant if $\|UAV\|_M = \|A\|_M$ for all A and all unitary U, V.

Table 1: A snippet of the AffectNet matrix

AffectNet	*IsA-pet*	*KindOf-food*	*Arises-joy*	...
dog	0.981	0	0.789	...
cupcake	0	0.922	0.910	...
songbird	0.672	0	0.862	...
gift	0	0	0.899	...
sandwich	0	0.853	0.768	...
rotten fish	0	0.459	0	...
win lottery	0	0	0.991	...
bunny	0.611	0.892	0.594	...
police man	0	0	0	...
cat	0.913	0	0.699	...
rattlesnake	0.432	0.235	0	...
...	...	...	...	...

is, the larger a concept's component in the e_0 direction is, the more affectively positive it is likely to be. Concepts with negative e_0 components, then, are likely to have negative affective valence.

Thus, by exploiting the information sharing property of SVD, concepts with the same affective valence are likely to have similar features – that is, concepts conveying the same emotion tend to fall near each other in AffectiveSpace. Concept similarity does not depend on their absolute positions in the vector space, but rather on the angle they make with the origin. For example, concepts such as `beautiful day`, `birthday party`, and `make someone happy` are found very close in direction in the vector space, while concepts like `feel guilty`, `be laid off`, and `shed tear` are found in a completely different direction (nearly opposite with respect to the centre of the space).

The problem with this kind of representation is that it is not scalable: when the number of concepts and of semantic features grows, the AffectNet matrix becomes too high-dimensional and too sparse for SVD to be computed (Balduzzi 2013). Although there has been a body of research on seeking for fast approximations of the SVD, the approximate methods are at most ≈ 5 times faster than the standard one (Menon and Elkan 2011), making it not attractive for real-world big data applications.

It has been conjectured that there might be simple but powerful meta-algorithms underlying neuronal learning (Lee et al. 2011). These meta-algorithms should be fast, scalable, effective, with few-to-no specific assumptions, and biologically plausible (Balduzzi 2013). Optimizing all the $\approx 10^{15}$connections through the last few million years' evolution is very unlikely (Balduzzi 2013). Alternatively, nature probably only optimizes the global connectivity (mainly the white matter), but leaves the other details to randomness (Balduzzi 2013). In order to cope with the ever-growing number of concepts and semantic features, thus, we replace SVD with random projection (RP) (Bingham and Mannila 2001), a data-oblivious method, to map the original high-dimensional data-set into a much lower-dimensional subspace by using a Gaussian $\mathrm{N}(0,1)$ matrix, while preserving the pair-wise distances with high probability. This theoretically solid and empirically verified statement follows Johnson and Lindenstrauss's (JL) Lemma (Balduzzi 2013). The JL Lemma states that with high probability, for all pairs of points $\mathrm{x}, \mathrm{y} \in \mathrm{X}$ simultaneously,

$$\sqrt{\frac{m}{d}} \parallel x - y \parallel_2 (1-\varepsilon) \leq \parallel \Phi x - \Phi y \parallel_2 \leq \tag{3}$$

$$\leq \sqrt{\frac{m}{d}} \parallel x - y \parallel_2 (1+\varepsilon), \tag{4}$$

where X is a set of vectors in Euclidean space, d is the original dimension of this Euclidean space, m is the dimension of the space we wish to reduce the data points to, ε is a tolerance parameter measuring to what extent is the maximum allowed distortion rate of the metric space, and Φ is a random matrix.

Structured random projection for making matrix multiplication much faster was introduced in (Sarlos 2006). Achlioptas (Achlioptas 2003) proposed *sparse random projection* to replace the Gaussian matrix with i.i.d. entries in

$$\phi_{ji} = \sqrt{s} \begin{cases} 1 & \text{with prob. } \frac{1}{2s} \\ 0 & \text{with prob.} 1 - \frac{1}{s} \\ -1 & \text{with prob.} \frac{1}{2s} \end{cases}, \tag{5}$$

where one can achieve a $\times 3$ speedup by setting $s = 3$, since only $\frac{1}{3}$ of the data need to be processed. However, since our input matrix is already too sparse, we avoid using sparse random projection.

When the number of features is much larger than the number of training samples ($d \gg n$), subsampled randomized Hadamard transform (SRHT) is preferred, as it behaves very much like Gaussian random matrices but accelerates the process from $\mathcal{O}(n\,d)$ to $\mathcal{O}(n\,log\,d)$ time (Lu et al. 2013). Following (Tropp 2011) (Lu et al. 2013), for $d = 2^p$ where p is any positive integer, a SRHT can be defined as:

$$\Phi = \sqrt{\frac{d}{m}} \mathrm{RHD} \tag{6}$$

where

• m is the number we want to subsample from d features randomly.

• R is a random $m \times d$ matrix. The rows of R are m uniform samples (without replacement) from the standard basis of $\mathbb{R}^d$.

• H$\in$ $\mathbb{R}^{d \times d}$ is a normalized Walsh-Hadamard matrix, which is defined recursively: $H_d = \begin{bmatrix} H_{d/2} & H_{d/2} \\ H_{d/2} & H_{d/2} \end{bmatrix}$ with $H_2 = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix}$.

• D is a $d \times d$ diagonal matrix and the diagonal elements are i.i.d. Rademacher random variables.

Our subsequent analysis only relies on the distances and angles between pairs of vectors (i.e. the Euclidean geometry information), and it is sufficient to set the projected space to be logarithmic in the size of the data (Ailon and Chazelle 2010) and apply SRHT. The result is a new vector space model, AffectiveSpace 2 (Fig. 1), which preserves the semantic and affective relatedness of common-sense concepts while being highly scalable.

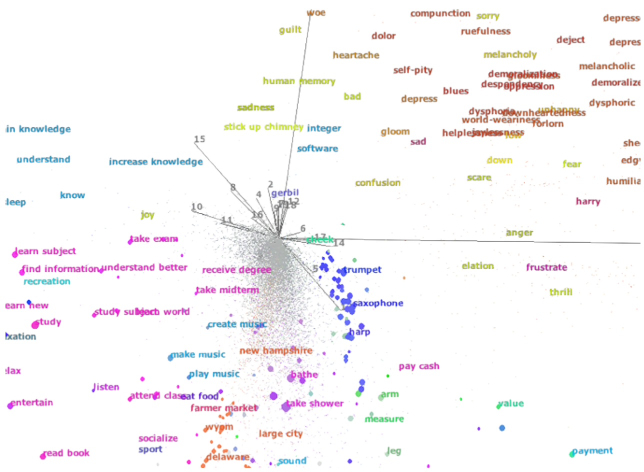

Figure 1: AffectiveSpace 2

Clustering AffectiveSpace 2

To reason on the disposition of concepts in AffectiveSpace 2, we use the Hourglass of Emotions (Fig. 2), an affective categorization model developed starting from Plutchik's studies on human emotions (Plutchik 2001). In the model, sentiments are re-organized around four independent dimensions whose different levels of activation make up the total emotional state of the mind. The Hourglass of Emotions, in fact, is based on the idea that the mind is made of different independent resources and that emotional states result from turning some set of these resources on and turning another set of them off (Minsky 2006).

In the model, affective states are not classified, as often happens in the field of emotion analysis, into basic emotional categories, but rather into four concomitant but independent dimensions, characterized by six levels of activation, which determine the intensity of the expressed/perceived emotion as a $float \in$ [-1,+1]. Such levels are also labeled as a set of 24 basic emotions (six for each of the affective dimensions) in a way that allows the model to specify the affective information associated with text both in a dimensional and in a discrete form.

Such basic emotions are used as initial centroids in AffectiveSpace 2 for clustering the vector space by means of sentic medoids (Cambria et al. 2011). Unlike the k-means algorithm (which does not pose constraints on centroids), sentic medoids do assume that centroids must coincide with k observed points, which allows to better cluster a vector space of common-sense knowledge. The sentic medoids approach is similar to the partitioning around medoids (PAM) algorithm, which determines a medoid for each cluster selecting the most centrally located centroid within that cluster. Unlike other PAM techniques, however, the sentic medoids algorithm runs similarly to k-means and, hence, requires a significantly reduced computational time. Generally, the initialization of clusters for clustering algorithms is a problematic task as the process often risks getting trapped in local optimum points, depending on the initial choice of centroids.

For this study, however, the set of 24 basic emotions of the Hourglass model are used as initial centroids. For this reason, what is usually seen as a limitation of the algorithm can be seen as advantage for this study, since what is being sought is not the k centroids leading to the best k clusters, but indeed the k centroids identifying the emotions we are interested in. Therefore, given that the distance between two points in the space is defined as $D(e_i, e_j) = \sqrt{\sum_{s=1}^{d'} \left(e_i^{(s)} - e_j^{(s)}\right)^2}$, the adopted algorithm can be summarized as follows:

1. Each centroid $\bar{e}_i \in \mathbb{R}^{d'}$ $(i = 1, 2, ..., k)$ is set as one of the 24 basic emotions of the Hourglass model;
2. Assign each instance e_j to a cluster $\bar{e}_i$ if $D(e_j, \bar{e}_i) \leq D(e_j, \bar{e}_{i'})$ where $i(i') = 1, 2, ..., k$;
3. Find a new centroid $\bar{e}_i$ for each cluster c so that $\sum_{j \in Cluster\ c} D(e_j, \bar{e}_i) \leq \sum_{j \in Cluster\ c} D(e_j, \bar{e}_{i'})$;
4. Repeat step 2 and 3 until no changes on centroids are observed.

Experimental Results

In order to evaluate the new analogical reasoning model, a comparison between AffectiveSpace and AffectiveSpace 2 has been performed both over a benchmark for affective common-sense knowledge (BACK) (Cambria and Hussain 2012b), for directly testing the affective analogical reasoning capabilities of the two models, and over a dataset of natural language opinions, for comparing how the two different configurations of AffectiveSpace (SVD-built versus RP-built) perform within the more practical task of concept-level opinion mining. Both vector space models were built upon the new 50k×120k AffectNet matrix.

Mood-Tag Evaluation

We compared AffectiveSpace and AffectiveSpace 2 on BACK, a benchmark for affective common-sense knowledge built by applying concept frequency - inverse opinion frequency (CF-IOF) (Cambria et al. 2010) on a 5,000-blogpost database extracted from LiveJournal[2], a virtual community of users who keep a blog, journal, or diary.

An interesting feature of this website is that bloggers are allowed to label their posts with both a category and a mood tag, by choosing from predefined categories and mood themes or by creating new ones. Since the indication of mood tags is optional, posts are likely to reflect the true mood of the authors.

CF-IOF weighting was exploited to filter out common concepts in the LiveJournal corpus and detect relevant mood-dependent semantics for each of the Hourglass sentic levels. The result was a benchmark of 2,000 affective concepts that were screened by 21 English-speaking students who were asked to evaluate the level b associated to each concept $b \in \Theta = \{\theta \in \mathbb{Z} \mid -1 \leq \theta \leq 1\}$ for each of the four affective dimensions. BACK's concepts were compared

[2]http://livejournal.com

with the classification results obtained by applying AffectiveSpace and AffectiveSpace 2, showing a consistent boost in classification performance (Table 2).

Sentic Computing Engine

The sentic computing engine (Cambria and Hussain 2012b) consists of four main components: a pre-processing module, which performs a first skim of text; a semantic parser, to deconstruct text into concepts; the IsaCore module, for aspect extraction; and the AffectiveSpace module, for polarity detection (Fig. 3).

Although similar in their structure, the last two modules are intrinsically different for the kind of knowledge they leverage on and for the task they fulfill. IsaCore (Cambria et al. 2014) is a semantic network of common knowledge (vocabulary knowledge collected from the Web), which focuses on the IsA relationship (e.g., Pablo Picasso-IsA-artist). AffectiveSpace is a vector space of affective common-sense knowledge (trivial knowledge that would not normally be found on the Web) leveraging on multiple relationships (e.g., LocatedAt, IsUsedFor, Arises, etc.). Hence, while the former exploits semantics to perform the task of aspect extraction, the latter uses sentics (i.e., affective information) to infer the polarity of natural language concepts.

Hourglass Interval	Sentic Level	AffSpace Accuracy	AffSpace 2 Accuracy
[G(1),G(2/3))	ecstasy	77.3%	84.5%
[G(2/3), G(1/3))	joy	83.9%	90.1%
[G(1/3),G(0))	serenity	68.8%	76.3%
(G(0), –G(1/3)]	pensiveness	74.5%	79.0%
(–G(1/3), –G(2/3)]	sadness	81.2%	89.6%
(–G(2/3), –G(1)]	grief	79.5%	87.4%

Table 2: Comparative evaluation of AffectiveSpace and AffectiveSpace 2 over the classification of Pleasantness sentic levels.

The engine does not aim to deeply understand natural language text, but rather to simply infer the denotative and connotative information associated with relevant concepts. In order to infer the polarity of a sentence, in fact, the sentic computing engine only needs to extract the features or aspects of the discussed service or product, e.g., size or weight of a phone, and the sentiments associated with each of these, e.g., positive or negative, so that the output of a sentence such as "I love the phone's touchscreen but its battery life is too short" would be something like <touchscreen: +> and <battery: –>.

The pre-processing module firstly exploits linguistic dictionaries to interpret all the affective valence indicators usually contained in opinionated text, e.g., special punctuation, complete upper-case words, cross-linguistic onomatopoeias, exclamation words, degree adverbs, and emoticons. Secondly, the module detects negation and spreads it in a way

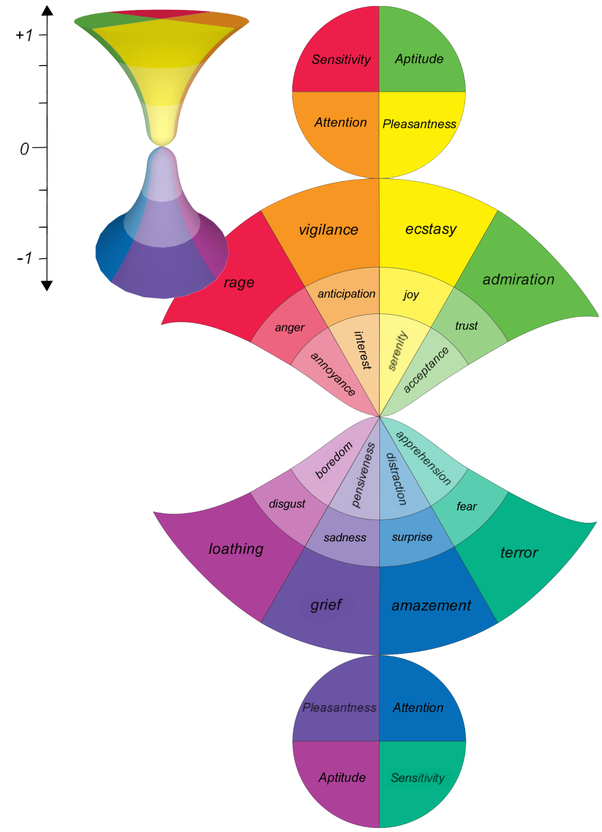

Figure 2: The Hourglass model

that it can be accordingly associated to concepts during the parsing phase. Such task is not trivial as not all appearances of explicit negation terms reverse the polarity of the enclosing sentence and that negation can often be expressed in rather subtle ways. Lastly, the module converts text to lower-case and, after lemmatizing it, splits the opinion into single clauses according to grammatical conjunctions.

For parsing text, the sentic parser is exploited for identifying concepts without requiring time-consuming phrase structure analysis. The parser uses knowledge about the lexical items found in text to choose the best possible construction for each span of text. Specifically, it looks each lexical item up in AffectNet and IsaCore, obtaining information about the basic category membership of that word. It then efficiently compares these potential memberships with the categories specified for each construction in the corpus, finding the best matches so that, for example, a concept like `buy christmas present` can be extracted from sentences such as "today I bought a lot of very nice Christmas gifts". Additionally, the sentic parser provides, for each retrieved concept, its relative frequency, valence, and status, i.e., the concept's occurrence in the text, its positive or negative connotation, and the degree of intensity with which the concept is expressed, respectively.

For each clause, the module outputs a small bag of concepts (SBoC), which is later on analyzed separately by the

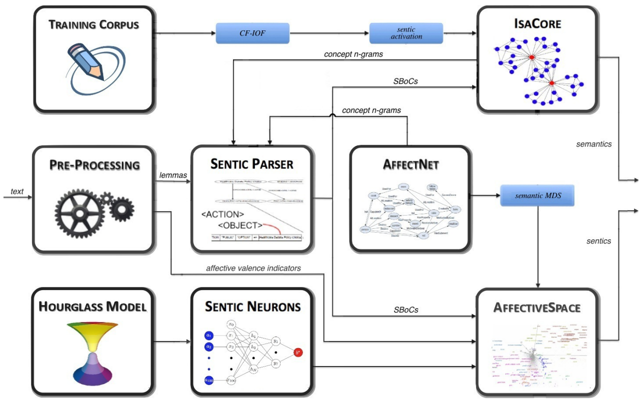

Figure 3: Block diagram of the sentic computing engine

IsaCore and AffectiveSpace modules. While the former exploits the graph representation of the common and common-sense knowledge base to detect semantics, the latter exploits the vector space representation of AffectNet to infer sentics. In particular, the IsaCore module applies spectral association for assigning activation to key nodes of the semantic network, which are used as seeds or centroids for classification. Such seeds can simply be the concepts corresponding to the class labels of interest or can be found by applying CF-IOF on a training corpus.

After seed concepts are identified, the module spreads their values across the IsaCore graph. This operation, an approximation of many steps of spreading activation, transfers the most activation to concepts that are connected to the seed concepts by short paths or many different paths in affective common-sense knowledge. Therefore, the concepts of each SBoC provided by the sentic parser are projected on the matrix resulting from spectral association in order to calculate their semantic relatedness to each seed concept and, hence, their degree of belonging to each different class. Such classification measure is directly proportional to the degree of connectivity between the nodes representing the retrieved concepts and the seed concepts in the IsaCore graph. The concepts retrieved by the sentic parser are also given as input to the AffectiveSpace module, which, in turn, exploits dimensionality reduction to infer the affective information associated with them. To this end, the concepts of each SBoC are projected into AffectiveSpace and, according to their position in the vector space, they are assigned to an affective class specified by the Hourglass model. In order to test the performance of the proposed model, such an operation is performed both with the SVD-built AffectiveSpace and with AffectiveSpace 2.

These are embedded in the sentic computing engine and evaluated against a dataset obtained from PatientOpinion[3], a social enterprise pioneering an online feedback service for users of the UK national health service. It is a manually tagged dataset of 2,000 patient opinions that associates to each post a category (namely, clinical service, communication, food, parking, staff, and timeliness) and a positive or negative polarity.

The dataset is hereby used to test the combined detection of opinion targets and the polarity associated with these. Results show that AffectiveSpace 2 generally outperforms standard AffectiveSpace, especially for categories where polarity is more difficult to detect in which affect is usually conveyed more implicitly, e.g., 'communication' and 'timeliness' (Table 3).

[3]http://patientopinion.org.uk

	AffSpace Accuracy	**AffSpace 2 Accuracy**
clinical service	75.2%	80.8%
communication	74.5%	85.1%
food	82.0%	83.7%
parking	74.0%	74.0%
staff	81.1%	83.2%
timeliness	73.4%	84.6%

Table 3: F-measure values relative to PatientOpinion evaluation.

Conclusions and Future Work

In a world in which millions of people express their opinions about commercial products and services everywhere on the Web, the distillation of knowledge from this huge amount of unstructured information is a key factor for tasks such as social media marketing, product positioning, and financial market prediction.

Common-sense reasoning is a good solution to the problem of concept-level sentiment analysis but scalability is a major factor in jeopardizing the efficiency of analogical reasoning in a multi-dimensional vector space of concepts.

In this work, we presented AffectiveSpace 2, a language visualization and analysis system that allows for reasoning by analogy on natural language concepts, even when these are represented by highly dimensional semantic features.

In this sense, AffectiveSpace 2 can be seen as a powerful tool for tackling the emerging issue of "Big Dimensionality" in the context of NLP, but also as a general framework for analogical reasoning that can be embedded in potentially any cognitive system dealing with real-world semantics.

References

Achlioptas, D. 2003. Database-friendly random projections: Johnson-lindenstrauss with binary coins. *Journal of computer and System Sciences* 66(4):671–687.

Ailon, N., and Chazelle, B. 2010. Faster dimension reduction. *Communications of the ACM* 53(2):97–104.

Balduzzi, D. 2013. Randomized co-training: from cortical neurons to machine learning and back again. *arXiv preprint arXiv:1310.6536*.

Bell, D.; Koulouri, T.; Lauria, S.; Macredie, R.; and Sutton, J. 2014. Microblogging as a mechanism for human-robot interaction. *Knowledge-Based Systems* 69:64–77.

Bingham, E., and Mannila, H. 2001. Random projection in dimensionality reduction: applications to image and text data. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, 245–250. ACM.

Cambria, E., and Hussain, A. 2012a. Sentic album: Content-, concept-, and context-based online personal photo management system. *Cognitive Computation* 4(4):477–496.

Cambria, E., and Hussain, A. 2012b. *Sentic Computing: Techniques, Tools, and Applications*. Dordrecht, Netherlands: Springer.

Cambria, E.; Hussain, A.; Durrani, T.; Havasi, C.; Eckl, C.; and Munro, J. 2010. Sentic computing for patient centered application. In *IEEE ICSP*, 1279–1282.

Cambria, E.; Mazzocco, T.; Hussain, A.; and Eckl, C. 2011. Sentic medoids: Organizing affective common sense knowledge in a multi-dimensional vector space. In Liu, D.; Zhang, H.; Polycarpou, M.; Alippi, C.; and He, H., eds., *Advances in Neural Networks*, volume 6677 of *Lecture Notes in Computer Science*, 601–610. Berlin: Springer-Verlag.

Cambria, E.; Song, Y.; Wang, H.; and Howard, N. 2014. Semantic multi-dimensional scaling for open-domain sentiment analysis. *IEEE Intelligent Systems* 29(2):44–51.

Fauconnier, G., and Turner, M. 2003. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. Basic Books.

Jolliffe, I. 2005. *Principal component analysis*. Wiley Online Library.

Jung, S., and Segev, A. 2014. Analyzing future communities in growing citation networks. *Knowledge-Based Systems* 69:34–44.

Lee, H.; Grosse, R.; Ranganath, R.; and Ng, A. Y. 2011. Unsupervised learning of hierarchical representations with convolutional deep belief networks. *Communications of the ACM* 54(10):95–103.

Li, X.; Xie, H.; Chen, L.; Wang, J.; and Deng, X. 2014. News impact on stock price return via sentiment analysis. *Knowledge-Based Systems* 69:14–23.

Lu, Y.; Dhillon, P.; Foster, D. P.; and Ungar, L. 2013. Faster ridge regression via the subsampled randomized hadamard transform. In *Advances in Neural Information Processing Systems*, 369–377.

Menon, A. K., and Elkan, C. 2011. Fast algorithms for approximating the singular value decomposition. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 5(2):13.

Minsky, M. 2006. *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York: Simon & Schuster.

Montejo-Raez, A.; Diaz-Galiano, M.; Martinez-Santiago, F.; and Urena-Lopez, A. 2014. Crowd explicit sentiment analysis. *Knowledge-Based Systems* 69:134–139.

Plutchik, R. 2001. The nature of emotions. *American Scientist* 89(4):344–350.

Poria, S.; Cambria, E.; Hussain, A.; and Huang, G.-B. 2015. Towards an intelligent framework for multimodal affective data analysis. *Neural Networks, DOI:10.1016/j.neunet.2014.10.005*.

Principi, E.; Squartini, S.; Cambria, E.; and Piazza, F. 2015. Acoustic template-matching for automatic emergency state detection: An elm based algorithm. *Neurocomputing* 149:426–434.

Rill, S.; Reinel, D.; Scheidt, J.; and Zicari, R. 2014. Politwi: Early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis. *Knowledge-Based Systems* 69:14–23.

Sarlos, T. 2006. Improved approximation algorithms for large matrices via random projections. In *Foundations of Computer Science, 2006. FOCS'06. 47th Annual IEEE Symposium on*, 143–152. IEEE.

Tropp, J. A. 2011. Improved analysis of the subsampled randomized hadamard transform. *Advances in Adaptive Data Analysis* 3(01n02):115–126.

Wang, Q.; Cambria, E.; Liu, C.; and Hussain, A. 2013. Common sense knowledge for handwritten chinese recognition. *Cognitive Computation* 5(2):234–242.

Zhai, Y.; Ong, Y.-S.; and Tsang, I. 2014. The emerging "big dimensionality". *IEEE Computational Intelligence Magazine* 9(3):14–26.