Semantic Outlier Detection for Affective Common-Sense Reasoning and Concept-Level Sentiment Analysis

Erik Cambria
School of Computer Engineering
Nanyang Technological University, Singapore
cambria@ntu.edu.sg

Giuseppe Melfi
Institut de l'Entreprise
Université de Neuchâtel, Switzerland
giuseppe.melfi@unine.ch

Abstract

Between the dawn of the Internet through year 2003, there were just a few dozens exabytes of information on the Web. Today, that much information is created weekly. The opportunity to capture the opinions of the general public about social events, political movements, company strategies, marketing campaigns, and product preferences has raised increasing interest both in the scientific community, for the exciting open challenges, and in the business world, for the remarkable fallouts in social media marketing and financial forecast. Keeping up with the ever-growing amount of unstructured information on the Web, however, is a formidable task. Unlike standard statistical approaches, sentic computing relies on a vector space model of affective common-sense knowledge to work with natural language at concept-level. The well-known noisiness of common-sense data sources, however, is a major factor in jeopardizing the efficiency of analogical reasoning in the vector space. In this work, it is explored how least absolute deviations can aid semantic outlier detection and, hence, enhance concept-level opinion mining.

Introduction

The ways people express their opinions and sentiments have radically changed in the past few years thanks to the advent of social networks, web communities, blogs, wikis, and other online collaborative media. Actually, the availability of these new tools allow people to create and share, in a time and cost efficient way, their own contents, ideas, and opinions virtually with the millions of people connected to the World Wide Web. This has made available by click a huge source of information and has provided a powerful communication medium to share knowledge and to get advantage from others' experiences.

As a major consequence, the distillation of knowledge from this huge amount of unstructured information can be a key factor for marketers who want to create an image or identity in the minds of their customers for their product, brand, or organization. On the other hand, these online social data remain hardly accessible to computers, as they are specifically meant for human consumption.

Online information retrieval is still mainly based on algorithms relying on the textual representation of web pages. Such algorithms are very good at retrieving texts, splitting them into parts, checking the spelling, and counting their words. But when it comes to interpreting sentences and extracting useful information for users, their capabilities are still very limited.

Indeed, such a scenario has led to the emerging fields of opinion mining and sentiment analysis (Pang and Lee 2008; Liu 2012; Cambria et al. 2013), which deal with information retrieval and knowledge discovery from text using data mining and natural language processing (NLP) techniques to distill knowledge and opinions from the huge amount of information on the World Wide Web. Mining opinions and sentiments from natural language, though, is an extremely difficult task as it involves a deep understanding of most of the explicit and implicit, regular and irregular, syntactical and semantic rules proper of a language.

Sentic computing (Cambria and Hussain 2015) tackles these crucial issues by exploiting affective common-sense reasoning, i.e., the intrinsically-human capacity to interpret the cognitive and affective information associated with natural language and, hence, to infer new knowledge and make decisions, in connection with one's social and emotional values, censors, and ideals. Thus, common-sense computing techniques are applied to bridge the semantic gap between word-level natural language data and the concept-level opinions conveyed by these.

To achieve this goal, the sentic computing framework takes advantage of a set of linguistic patterns (Poria et al. 2014) and AffectNet[1], a semantic network in which common-sense concepts (e.g., 'read book', 'payment', 'play music') are linked to a hierarchy of affective domain labels (e.g., 'joy', 'amazement', 'fear', 'admiration'). In particular, the vector space representation of such a semantic network, termed AffectiveSpace[2] (Cambria et al. 2015a), enables affective analogical reasoning on natural language concepts. In practice, concepts conveying similar semantic and affective information, e.g., 'enjoy conversation' and 'chat with friend', tend to fall near each other in the multi-dimensional space that characterizes AffectiveSpace (Fig. 1).

[1] `http://sentic.net/affectnet.zip`
[2] `http://sentic.net/affectivespace.zip`

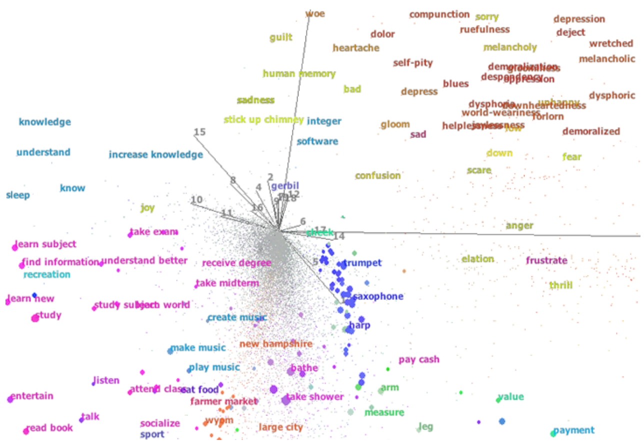

Figure 1: AffectiveSpace

One major issue in performing reasoning in AffectiveSpace is the noisy nature of the common-sense data the vector space is built upon. Wrong semantic and affective features associated with natural language concepts heavily affect the configuration of the vector space and the clustering performance. To this end, the present research work aims to explore how least absolute deviations (LAD) can aid semantic outlier detection and, hence, enhance analogical reasoning in AffectiveSpace.

The rest of the paper is organized as follows: next section presents related work in the field of concept-level sentiment analysis; the following two sections describe the multidimensional vector space model of affective common-sense knowledge and the adopted emotion model, respectively; next follows a section showing how and why semantic outlier detection is performed in the vector space model; finally, the paper is concluded by an evaluation section and a final section offering closing remarks.

Background

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 financial market prediction. The potential applications of concept-level sentiment analysis, in fact, are countless and span interdisciplinary areas such as political forecasting, brand positioning, 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.

The Vector Space Model

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. Marvin Minsky attributes this property to the so called 'difference-engines' (Minsky 1986).

This particular kind of agents operates by recognizing differences between the current state and the desired state, and acting to reduce each difference by invoking K-lines that turn on suitable solution methods. This kind of thinking is maybe the essence of our supreme intelligence since in everyday life no two situations are ever the same and have to perform this action continuously. 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.

To this end, principal component analysis (PCA) has been applied on the matrix representation of AffectNet. In particular, truncated singular value decomposition (TSVD) has been preferred to other dimensionality reduction techniques for its simplicity, relatively low computational cost, and compactness. TSVD, in fact, is particularly suitable for measuring the cross-correlations between affective common-sense concepts as it uses an orthogonal transformation to convert the set of possibly correlated common-sense features associated with each concept into a set of values of uncorrelated variables (the principal components of the SVD). By using Lanczos' method (Lanczos 1950), moreover, the generalization process is relatively fast (a few seconds), despite the size and the sparseness of AffectNet. Applying TSVD on AffectNet causes it to describe other features that could apply to known affective concepts by analogy: if a concept in the matrix has no value specified for a feature owned by many similar concepts, then by analogy the concept is likely to have that feature as well.

In other words, concepts and features that point in similar directions and, therefore, have high dot products, are good candidates for analogies. After performing TSVD on AffectNet, hereby termed A for the sake of conciseness, a low-rank approximation of it is obtained, that is, a new matrix $\tilde{A} = U_p \Sigma_p V_p^T$. This approximation is based on minimizing the Frobenius norm of the difference between A and $\tilde{A}$ under the constraint $rank(\tilde{A}) = p$. For the Eckart–Young theorem (Eckart and Young 1936), it represents the best approximation of A in the least-square sense, in fact:

$$\begin{aligned}\min_{\tilde{A}|rank(\tilde{A})=p} |A-\tilde{A}| &= \min_{\tilde{A}|rank(\tilde{A})=p} |\Sigma - U_p^T \tilde{A} V_p| \\ &= \min_{\tilde{A}|rank(\tilde{A})=p} |\Sigma - S_p|\end{aligned}$$

assuming that $\tilde{A}$ has the form $\tilde{A} = USV^*$, where S is diagonal.

From the rank constraint, i.e., S has p non-zero diagonal entries, the minimum of the above statement is obtained as follows:

$$\min_{\tilde{A}|rank(\tilde{A})=p} |\Sigma - S_p| = \min_{s_i} \sqrt{\sum_{i=1}^{n} (\sigma_i - s_i)^2} =$$

$$= \min_{s_i} \sqrt{\sum_{i=1}^{p} (\sigma_i - s_i)^2 + \sum_{i=p+1}^{n} \sigma_i^2} = \sqrt{\sum_{i=p+1}^{n} \sigma_i^2}$$

Therefore, $\tilde{A}$ of rank p is the best approximation of A in the Frobenius norm sense when $\sigma_i = s_i$ $(i = 1, ..., p)$ and the corresponding singular vectors are the same as those of A. If all but the first p principal components are discarded, common-sense concepts and emotions are represented by vectors of p 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_{p-1} of the vector space. For example, the most significant eigenmood, e_0, represents concepts with positive affective valence.

That 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 TSVD, 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 key to perform common-sense reasoning is to find a good trade-off for representing knowledge. Since in life two situations are never exactly the same, no representation should be too concrete, or it will not apply to new situations, but, at the same time, no representation should be too abstract, or it will suppress too many details. Within AffectiveSpace, this knowledge representation trade-off can be seen in the choice of the vector space dimensionality. The number p of singular values selected to build AffectiveSpace is a measure of the trade-off between precision and efficiency in the representation of the affective common-sense knowledge base.

The bigger is p, the more precisely AffectiveSpace represents AffectNet's knowledge, but generating the vector space is slower, and so is computing dot products between concepts. The smaller is p, on the other hand, the more efficiently AffectiveSpace represents affective common-sense knowledge both in terms of vector space generation and of dot product computation. However, too few dimensions risk not to correctly represent AffectNet as concepts defined with too few features tend to be too close to each other in the vector space and, hence, not easily distinguishable and clusterable.

The Emotion Model

The Hourglass of Emotions (Cambria, Livingstone, and Hussain 2012) is an affective categorization model, inspired by Plutchik's studies on human emotions (Plutchik 2001), which organizes primary affective states around four independent but concomitant dimensions, whose different levels of activation make up the total emotional state of the mind. Such a reinterpretation is inspired by Minsky's theory of the mind, according to which brain activity consists 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).

Each affective dimension of the Hourglass model is characterized by six levels of activation (measuring the strength of an emotion), termed 'sentic levels', which represent the intensity thresholds of the expressed or perceived emotion. These 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.

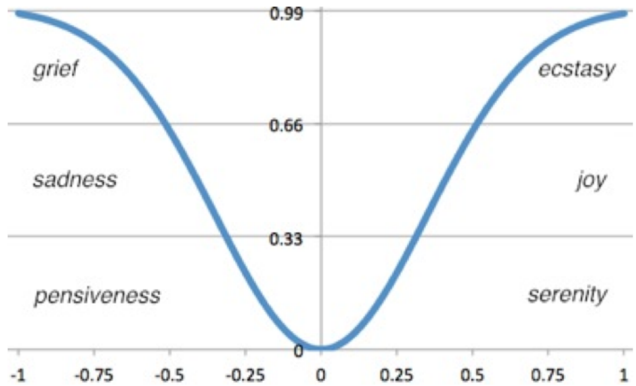

Figure 2: Pleasantness emotional flow

The transition between different emotional states is modeled, within the same affective dimension, using the function $G(x) = 1 - \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2 / 2\sigma^2}$, for its symmetric inverted bell curve shape that quickly rises up towards the unit value (Fig. 2). Beyond emotion detection, the Hourglass model is also used for polarity detection tasks. Since polarity is strongly connected to attitudes and feelings, it is defined in terms of the four affective dimensions, according to the formula:

$$p = \sum_{i=1}^{N} \frac{Plsn(c_i) + |Attn(c_i)| - |Snst(c_i)| + Apti(c_i)}{3N}$$

where c_i is an input concept, N the total number of concepts, and 3 the normalization factor (as the Hourglass dimensions are defined as float $\in$ [-1,+1]). In the formula, Attention is taken as absolute value since both its positive and negative intensity values correspond to positive polarity values (e.g., 'surprise' is negative in the sense of lack of Attention, but positive from a polarity point of view). Similarly, Sensitivity is taken as negative absolute value since both its positive and negative intensity values correspond to negative polarity values (e.g., 'anger' is positive in the sense of level of activation of Sensitivity, but negative in terms of polarity).

Outlier-Free Reasoning

Affective analogical reasoning consists in processing the cognitive and affective information associated with natural language concepts, in order to compare the similarities between new and understood concepts and, hence, use such similarities to gain understanding of the new concept. It is a form of inductive reasoning because it strives to provide understanding of what is likely to be true, rather than deductively proving something as fact.

The reasoning process begins by determining the target concept to be learned or explained. It is then compared to a general matching concept whose semantics and sentics (that is, the conceptual and affective information associated with it) are already well-understood. The two concepts must be similar enough to make a valid, substantial comparison.

Affective analogical reasoning is based on the brain's ability to form semantic patterns by association. The brain may be able to understand new concepts more easily if they are perceived as being part of a semantic pattern. If a new concept is compared to something the brain already knows, it may be more likely that the brain will store the new information more readily. Clearly, such a semantic association heavily depends on semantic features shared by the represented concepts. Concept similarity, in fact, does not depend on their absolute positions in the vector space, but rather on the angle they make with the origin.

For this reason, the presence of semantic outliers, that is, concepts that share a few semantic features with a group of neighbor concepts without in fact being affectively-related, can heavily affect the configuration of AffectiveSpace and, hence, affective analogical reasoning. In order to tackle such a crucial issue, least absolute deviations (LAD) regression is applied.

Least Absolute Deviations

The robustness of LAD to low-leverage outliers, and its susceptibility to high-leverage outliers has been extensively studied in literature (Dodge 1987; 1997; 2002). This paper adopts a method for non-parametic detection of such influential observations by the use of a technique derived from LAD regression (Faria and Melfi 2006). Let $S \subset \mathbb{R}^{p+1}$ be a finite discrete set of points, represented by $p+1$ variables. Denote the elements of S as $(x_{i1}, \ldots, x_{ip}, y_i)$, where the last variable is explained from the preceding ones by a linear regression model:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i \qquad \text{for } i = 1, \ldots, n$$

where p is the number of explanatory variables, ε_i are error terms, or deviations, and n is the number of observations. The LAD regression model is determined by minimizing the sum of the absolute deviations, i.e., the vector $(\beta_0, \beta_1, \ldots, \beta_p) \in \mathbb{R}^{p+1}$ is determined by minimizing on $\beta_0, \beta_1, \ldots, \beta_p$ the function

$$F(\beta_0, \beta_1, \ldots, \beta_p) = \sum_{i=1}^{n} \left| y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right|$$

When a linear LAD regression model is fitted, the hyperplane always passes through at least $p+1$ points (Arthanari and Dodge 1993), although the solution may be non-unique. For our purposes, it is assumed that the hyperplane which fits the linear LAD regression model is unique, as well as the observation with maximal absolute deviation.

Semantic Outlier Detection

Applying LAD regression to AffectiveSpace when clustering it according to each different affective dimension, allows concepts that are not affectively-related to the sentic levels of the Hourglass model to be excluded from the affective analogical reasoning process. In particular, AffectiveSpace is clustered four times, in order for each commonsense concept to be expressed by a four-dimensional vector, which synthesizes the affective information conveyed by it in terms of Pleasantness, Attention, Sensitivity, and Aptitude. For the Pleasantness dimension, for example, the analogical reasoning process aims to detect concepts conveying ‘ecstasy’, ‘joy’, ‘serenity’, ‘pensiveness’, ‘sadness’, and ‘grief’, respectively. Because of how AffectiveSpace is designed, such an operation often includes common-sense concepts that, despite carrying similar semantics, are not affectively-related, that is, do not convey any of those emotions or derivates. Examples include concepts like ‘newspaper’, which is found for its semantic-relatedness to ‘give good news’, but which does not convey any emotion, or even ‘make fun of someone’, which is detected in its capacity as semantic relative of ‘make someone happy’, but which instead carries emotions that belong to a different affective dimension.

Experimental results

In order to evaluate the proposed semantic outlier detection technique, a comparison between an outlier-contaminated and an outlier-free 100-dimensional AffectiveSpace has been performed both over a benchmark for affective common-sense knowledge (BACK) (Cambria and Hussain 2015), for directly testing the affective analogical reasoning capabilities of the two frameworks, and over a dataset of natural language opinions, for comparing how the two different configurations of AffectiveSpace (standard versus outlier-free) perform within the more practical task of concept-level opinion mining.

Mood-Tag Evaluation

We compared standard AffectiveSpace and outlier-free AffectiveSpace on BACK, a benchmark for affective common-sense knowledge built by applying concept frequency - inverse opinion frequency (CF-IOF) (Cambria and Hussain 2015) on a 5,000-blogpost database extracted from LiveJournal[3], 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 $b \in \Theta = \{-1, -2/3, -1/3, 0, 1/3, 2/3, 1\} = \{\theta/3 \mid \theta \in \mathbb{Z}, -1 \leq \theta/3 \leq 1\}$ for each of the four affective dimensions. Results obtained were averaged. BACK’s concepts were compared with the classification results obtained by applying the standard AffectiveSpace process and outlier-free AffectiveSpace, showing a consistent boost in classification performance (Table 1).

[3]http://livejournal.com

Hourglass Interval	Sentic Level	Standard AffSpace	Outlier-Free AffSpace
[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 1: Comparative evaluation of standard and outlier-free AffectiveSpace over the classification of Pleasantness levels.

Sentic Computing Engine

The sentic computing engine consists of three main components: a pre-processing module, which performs a first skim of the opinion; a semantic parser, whose aim is to extract concepts from the opinionated text; the AffectiveSpace module, for inferring the semantics and sentics associated with the given concepts. 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 opinion target aspects and the sentiments associated with each of these.

The pre-processing module 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.

The semantic 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, 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” (Cambria et al. 2015b).

The concepts retrieved by the semantic parser 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 semantic outlier detection technique, such an operation is performed both with standard AffectiveSpace and with outlier-free AffectiveSpace. These are embedded in the sentic computing engine and evaluated against a dataset obtained from PatientOpinion[4], a manually tagged dataset of 2,000 patient opinions that associates to each post a category and a positive or negative polarity.

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

	Standard AffectiveSpace	Outlier-Free AffectiveSpace
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 2: F-measure values relative to PatientOpinion evaluation.

The dataset is hereby used to test the combined detection of opinion targets and the polarity associated with these. Results show that outlier-free AffectiveSpace 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 2).

Conclusion

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 the well-known noisiness of common-sense data sources is a major factor in jeopardizing the efficiency of analogical reasoning in a multi-dimensional vector space of concepts. In this work, we showed how least absolute deviations can aid semantic outlier detection and, hence, consistently enhance analogical reasoning for tasks such as emotion recognition and aspect-based opinion mining.

References

Arthanari, T., and Dodge, Y. 1993. *Mathematical Programming in Statistics*. New York: John Wiley and Sons, classic edition.

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.

Cambria, E., and Hussain, A. 2015. *Sentic Computing: A Common-Sense-Based Framework for Concept-Level Sentiment Analysis*. Cham, Switzerland: Springer.

Cambria, E.; Schuller, B.; Liu, B.; Wang, H.; and Havasi, C. 2013. Knowledge-based approaches to concept-level sentiment analysis. *IEEE Intelligent Systems* 28(2):12–14.

Cambria, E.; Fu, J.; Bisio, F.; and Poria, S. 2015a. AffectiveSpace 2: Enabling affective intuition for concept-level sentiment analysis. In *AAAI*.

Cambria, E.; Gastaldo, P.; Bisio, F.; and Zunino, R. 2015b. An ELM-based model for affective analogical reasoning. *Neurocomputing* 149:443–455.

Cambria, E.; Livingstone, A.; and Hussain, A. 2012. The hourglass of emotions. In Esposito, A.; Vinciarelli, A.; Hoffmann, R.; and Muller, V., eds., *Cognitive Behavioral Systems*, volume 7403 of *Lecture Notes in Computer Science*. Berlin Heidelberg: Springer. 144–157.

Dodge, Y., ed. 1987. *Statistical data analysis based on the L1-norm and related methods*. New York; Amsterdam: Elsevier/North-Holland.

Dodge, Y., ed. 1997. *L1-statistical procedures and related topics*. Institute of Mathematical Statistics.

Dodge, Y., ed. 2002. *Statistical Data Analysis Based on the L1-Norm and Related Methods*. Birkhäuser: Springer.

Eckart, C., and Young, G. 1936. The approximation of one matrix by another of lower rank. *Psychometrika* 1(3):211–218.

Faria, S., and Melfi, G. 2006. LAD regression and nonparametric methods for detecting outliers and leverage points. *Student* 5(3-4):265–272.

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

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

Lanczos, C. 1950. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of The National Bureau of Standards* 45(4):255–282.

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.

Liu, B. 2012. *Sentiment Analysis and Opinion Mining*. Morgan and Claypool.

Minsky, M. 1986. *The Society of Mind*. New York: Simon and Schuster.

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.

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

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

Poria, S.; Cambria, E.; Winterstein, G.; and Huang, G.-B. 2014. Sentic patterns: Dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems* 69:45–63.

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.