

Sentic Computing

Erik Cambria¹ · Amir Hussain²

Published online: 19 March 2015
© Springer Science+Business Media New York 2015

Guest Editor Introduction

As the Web rapidly evolves, Web users are evolving with it. In an era of social connectedness, people are becoming more and more enthusiastic about interacting, sharing, and collaborating through social networks, online communities, blogs, Wikis, and other online collaborative media. In recent years, this collective intelligence has spread on many different areas, with particular focus on fields related to everyday life such as commerce, tourism, education, and health, causing the size of the Social Web to expand exponentially.

The distillation of knowledge from such a large amount of unstructured information, however, is an extremely difficult task, as the contents of today's Web are perfectly suitable for human consumption, but remain hardly accessible to machines. The opportunity to capture the opinions of the general public about social events, political movements, company strategies, marketing campaigns, and product preferences has 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.

Mining opinions and sentiments from natural language, however, 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. Existing approaches mainly rely on parts of text, in which opinions and sentiments that are explicitly expressed, such as polarity terms, affect words and their co-occurrence frequencies. However, opinions and sentiments are often conveyed implicitly through latent semantics, which make purely syntactical approaches ineffective.

Sentic computing is a multi-disciplinary approach to sentiment analysis at the crossroads between affective computing and common-sense computing, which exploits both computer and social sciences to better recognize, interpret, and process opinions and sentiments over the Web. In sentic computing, whose term derives from the Latin *sentire* (root of words such as sentiment and sentience) and *sensus* (intended both as capability of feeling and as common sense), the analysis of natural language is based on affective ontologies and common-sense reasoning tools, which enable the analysis of text not only at document, page, or paragraph levels, but also at sentence, clause, and concept levels.

In particular, sentic computing involves the use of AI and Semantic Web techniques, for knowledge representation and inference; mathematics, for carrying out tasks such as graph mining and multi-dimensionality reduction; linguistics, for discourse analysis and pragmatics; psychology, for cognitive and affective modeling; sociology, for understanding social network dynamics and social influence; finally ethics, for understanding related issues about the nature of mind and the creation of emotional machines.

In this light, this Special Issue focuses on the introduction, presentation, and discussion of novel approaches that further develop and apply sentic computing models, tools, and techniques for the design of emotion-sensitive systems spanning areas such as computational intelligence,

✉ Erik Cambria
cambria@ntu.edu.sg

Amir Hussain
ahu@cs.stir.ac.uk

¹ Nanyang Technological University, Singapore, Singapore

² University of Stirling, Stirling, Scotland, UK

Semantic Web, fuzzy logic, linguistics, and sociology. Accordingly, the contained articles in this issue cover variegated topics including polarity detection, aspect extraction, semantic networks, emotion classification, semantic compositionality, transfer learning, big data, conceptual frames, and behavioral analytics. Out of the sixteen submissions received for this special issue, six were accepted. Four of the accepted papers underwent two rounds of revisions, and the rest underwent only one.

The article “Propagating and Aggregating Fuzzy Polarities for Concept-Level Sentiment Analysis” by Mauro Dragoni, Andrea Tettamanzi, and Celia da Costa Pereira investigates how sentiment polarities associated with concepts have to be adapted with respect to the different domains in which they are used. In particular, authors explore the use of fuzzy logic for modeling concept polarities, and the uncertainty associated with them, with respect to different domains. The approach is based on the use of a knowledge graph built by combining two linguistic resources, namely WordNet and SenticNet. Such a knowledge graph is then exploited by a graph-propagation algorithm that propagates sentiment information learned from labeled datasets. The system implementing the proposed approach has been evaluated on the Blitzer dataset. The results demonstrate its viability in real-world cases.

In “Resonance Elicits Diffusion: Modeling Subjectivity for Retweeting Behavior Analysis”, Songxian Xie, Jintao Tang, and Ting Wang explore the subjective motivations of retweeting, the core mechanism of information diffusion on Twitter. In light of psychological theory, authors assume that a tweet is more likely to be retweeted by a user because of similar subjectivity, and propose a subjectivity model to combine both the topics and opinions to model subjectivity. With state-of-the-art topic model and sentiment analysis techniques, authors design a subjectivity model by finding topics and determining opinions toward these topics from user-generated content simultaneously. Authors evaluate their model in the retweeting analysis problem to verify its impact on retweeting and effectiveness in the retweeting prediction performance.

Next, “Sentilo: Frame-based Sentiment Analysis” by Diego Reforgiato Recupero, Valentina Presutti, Sergio Consoli, Aldo Gangemi, and Andrea Giovanni Nuzzolese presents Sentilo, an unsupervised, domain-independent system that performs sentiment analysis by hybridizing natural language processing techniques and Semantic Web technologies. Given a sentence expressing an opinion, Sentilo recognizes its holder, detects the topics and subtopics that it targets, links them to relevant situations and events referred to by it, and evaluates the sentiment expressed on each topic/subtopic. Sentilo relies on a novel lexical resource, which enables a proper propagation of sentiment scores from topics to subtopics, and on a formal

model expressing the semantics of opinion sentences. Sentilo provides its output as a RDF graph, and whenever possible it resolves holders’ and topics’ identity on Linked Data.

The article “Word Embedding Composition for Data Imbalances in Sentiment and Emotion Classification” is presented by Ruifeng Xu, Tao Chen, Yunqing Xia, Qin Lu, Bin Liu, and Xuan Wang, who describe an oversampling method based on word embedding compositionality which produces meaningful balanced training data. Authors first use a large corpus to train a continuous skip-gram model to form a word embedding model maintaining the syntactic and semantic integrity of the word features. Then, a compositional algorithm based on recursive neural tensor networks is used to construct sentence vectors based on the word embedding model. Finally, authors use the SMOTE algorithm as an oversampling method to generate samples for the minority classes and produce a fully balanced training set. Evaluation results on two quite different tasks show that the feature composition method and the oversampling method are both important in obtaining improved classification results.

Following, “AspNet: Aspect Extraction by Bootstrapping Generalization and Propagation using an Aspect Network”—paper handled independently during review process—is elaborated upon by Yunqing Xia, Erik Cambria, and Amir Hussain, who introduce a novel unsupervised method to extract aspects from raw reviews with a broader coverage. Previous research has shown that unsupervised methods based on dependency relations are promising for opinion target extraction (OTE). In this work, authors introduce AspNet, an aspect network that further improves existing OTE methods by providing a new framework for modeling aspects. AspNet represents the general indecomposable atom aspects and their dependency relations in a two-layered, directed, weighted graph, based on which the specific decomposable compound aspects in reviews can be effectively extracted. AspNet is constructed through an unsupervised learning method that starts from a small number of human-defined, domain-dependent aspects, and bootstraps generalization and propagation in a large volume of raw reviews.

Finally, the contribution “Twitter Sentiment Analysis for Large Scale Data: An Unsupervised Approach” by Rafeeqe Pandarachalil, Selvaraju Sendhilkumar, and GS Mahalakshmi explores how unsupervised method for analyzing tweet sentiments can address the problem of scalability when dealing with an ever-growing amount of tweets. Polarity of tweets are evaluated by using three sentiment lexicons—SenticNet, SentiWordNet, and SentslangNet. SentslangNet is a sentiment lexicon built from SenticNet and SentiWordNet for slangs and acronyms. Experimental results show fairly good F score. The method

is implemented and tested in parallel python framework and is shown to scale well with large volume of data on multiple cores.

The guest editors are grateful to the 30 reviewers for their timely and insightful reviews of the submissions: Schuller Björn, Qiao Cai, Alexandra Balahur, Carlos Travieso González, Aitzol Ezeiza, Sascha Fagel, Radim

Burget, Igor Aleksander, Amine Chohra, Maiga Chang, Enzo Emanuele, Julian Szymanski, Vanda Zammuner, Emanuele Principi, Arshad Ali, Dorian Aur, Oya Aran, David Balduzzi, Bin Hu, Danilo Comminiello, Andre Fonseca, Anna Belardinelli, Polychronis Koutsakis, Vishal Goyal, Ricardo Ribeiro, Rui Xia, Rajiv Bajpai, Soujanya Poria, Dheeraj Rajagopal, and Iti Chaturvedi.