Multi-lingual support for lexicon-based sentiment analysis guided by semantics

Alexander Hogenboom, Bas Heerschop, Flavius Frasincar, Uzay Kaymak, Franciska de Jong

A Eramsus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands
B Eindhoven University of Technology, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands
C Universiteit Twente, P.O. Box 217, NL-7500 AE Enschede, The Netherlands

1 Introduction

In today’s complex, globalizing markets, information monitoring tools are of paramount importance for decision makers. Such tools help decision makers in identifying issues and patterns that matter, as well as in tracking and predicting emerging events. Traditional decision support systems typically provide support for decisions by accurately deriving actionable knowledge from structured data, whereas the extraction of useful information from unstructured data like natural language text still poses important challenges [1]. Recent advances in tools for information extraction have been primarily focused on retrieving explicit pieces of information from natural language text on different levels of granularity [2]. State-of-the-art information monitoring and extraction tools enable us to identify entities like companies, products, or brands in text, and to subsequently extract more complex concepts, such as events in which these entities play various roles [3]. Recent research endeavors additionally explore how to perform such information extraction tasks on a multitude of heterogeneous sources in an ever-changing environment [4–6].

However, latent pieces of information can be extracted from natural language text as well. For instance, recent work has made it possible to detect the distinct topics that people discuss in their (on-line) conversations [7,8]. Yet, for many application scenarios, it is not so much the entities, events, or topics that people discuss per se, but rather people’s sentiment with respect to these subjects that provides decision makers with valuable information. This is reflected by the recent surge in research interest in sentiment analysis for decision support [1,9–11].

Sentiment analysis techniques can support decision making in a multitude of scenarios. For instance, sentiment analysis can help organizations pinpoint the effect of specific issues on customer perceptions, thus helping these organizations respond with appropriate marketing and public relations strategies [12]. Furthermore, consumer sentiment has been demonstrated to have a significant impact on stock ratings [13,14] and sales [15,16]. Thus, accurate sentiment analysis methods are crucial for supporting decision making in these fields. Additionally, tracking of stakeholders’ sentiment is important for decision making in economic systems [17], financial markets [18], politics [19], organizations [20], and reputation management [21].

Real-world decision support systems typically consist of four logical components, i.e., a Knowledge Management System (KMS), a Model Management System (MMS), a Database Management System (DMS), and a User Interface System (UIS) [22]. Each of these logical components

Available online 20 March 2014
Accepted 11 March 2014
Received in revised form 11 February 2014
Received 24 September 2013
Article history:

© 2014 Elsevier B.V. All rights reserved.

Keywords:
Multi-lingual sentiment analysis
Sematics
Lexicon
Machine translation
Map
Propagation

Corresponding author. Tel.: +31 10 408 1340; fax: +31 10 408 9162.
E-mail addresses: hogenboom@ese.eur.nl (A. Hogenboom), basheerschop@gmail.com (B. Heerschop), frasincar@ese.eur.nl (F. Frasincar), ukaymak@ieee.org (U. Kaymak), fm.g.dejong@utwente.nl (F. de Jong).

http://dx.doi.org/10.1016/j.dss.2014.03.004
0167-9236/© 2014 Elsevier B.V. All rights reserved.
is used to monitor and regulate the flow of crucial information in order to support decision making in an organization. Data managed by the DMS can be transformed into actionable knowledge in the KMS, with the MMS controlling how the obtained knowledge is used in models in order to support decision making, and the UIS taking care of the interaction with the end user of the system. In order to utilize sentiment-based information in decision support systems, the DMS should be enriched with (user-generated) sentiment-carrying content that has been crawled from the Web. Furthermore, the KMS should be able to represent the indicators of identified sentiment with respect to a topic of interest. Additionally, the MMS should allow for the incorporation of sentiment-based information in the decision making process. Last, the UIS should provide dashboards with relevant information that enables decision makers to act upon arising issues in a timely manner.

One of the key open issues that must be resolved in order to be able to exploit the full potential of sentiment analysis in real-life decision support systems is that these systems must be able to deal with textual data in various languages [1]. Such data is available in vast amounts, as recent developments on the Web enable users to produce an ever-growing amount of virtual utterances of opinions or sentiment through, e.g., messages on Twitter, blogs, or reviews, in any language of their preference.

The analysis of sentiment in the overwhelming amount of available multi-lingual textual data is challenging at best. This challenge can be addressed by means of automated sentiment analysis techniques, focusing on determining the polarity of natural language text. Typical approaches involve scanning a text for cues signaling its polarity, e.g., (parts of) words or other (latent) features of natural language text. Lexicon-based sentiment analysis methods have gained (renewed) attention in recent work [23–29], not in the least because their performance has been shown to be robust across domains and texts [30]. Such methods essentially rely on lexical resources containing words and their associated sentiment, i.e., sentiment lexicons, and their nature allows for intuitive ways of accounting for structural or semantic aspects of text in sentiment analysis [26,31].

Many existing lexicon-based sentiment analysis approaches are tailored to one specific language — typically English. However, in order for automated sentiment analysis to be useful for decision makers in today’s complex, globalizing markets, automated sentiment analysis tools need to be able to support multiple languages rather than English only. Therefore, we explore how we can analyze sentiment in another language — i.e., Dutch — for which we have nothing more but some lexical and syntactical parsing tools, a semantic lexical resource, and a handful of positive and negative sample words.

A good starting point is SentiWordNet [32,33], as recent research has proven this large (semantic) sentiment lexicon for English, generated by means of machine learning techniques, to be rather effective when used for analyzing sentiment in texts published in our reference language, i.e., English [34]. As a first step, one could consider translating texts from a target language, i.e., Dutch, to our reference language, i.e., English, in order to be able to subsequently utilize the well-established SentiWordNet sentiment lexicon for the reference language in the sentiment analysis process.

However, as subjectivity is associated with word meanings rather than words [35], the literal translation of texts to a reference language in order to benefit from the available sentiment lexicon for the reference language may be suboptimal in automated sentiment analysis of texts in another language. As an alternative, we therefore propose to map the sentiment from the reference sentiment lexicon to a sentiment lexicon for the target language, by means of traversing relations between large language-specific semantic lexical resources, thus accounting for word meanings rather than lexical representations. Additionally, we consider an approach that involves propagating sentiment from a seed set of words in a language-specific semantic lexical resource for each considered language separately, in order to generate language-specific sentiment lexicons which can subsequently be used in language-specific sentiment analysis methods.

The main contribution of our work lies in our novel sentiment mapping method, which exploits relations between language-specific semantic lexicons in order to construct a sentiment lexicon for a target language. We compare the effectiveness of this method with that of an existing machine-translation approach and a method that focuses on semantic relations within, rather than across languages. We thus aim to provide insight in the importance of semantics for multi-lingual sentiment analysis.

The remainder of the paper is organized as follows. In Section 2, we discuss related work on (multi-lingual) sentiment analysis and the semantic lexicons that may be exploited in this process. We then elaborate on our framework for assessing our considered methods for dealing with another language in sentiment analysis in Section 3. Our findings are discussed in Section 4. We conclude and provide directions for future work in Section 5.

2. Related work

Today’s abundance of user-generated content has resulted in a surge of research interest in systems that are able to deal with opinions and sentiment, as explicit information on user opinions is often hard to find, confusing, or overwhelming [36]. Many language-specific sentiment analysis approaches exist, whereas the exploration of how to support multiple languages when analyzing sentiment has only just begun.

2.1. Sentiment analysis

The roots of sentiment analysis are in fields like natural language processing, computational linguistics, and text mining. The main objective of most sentiment analysis approaches is to extract subjective information from natural language text. Most work focuses on determining the overall polarity of words, sentences, text segments, or documents [36]. This task is commonly approached as a binary classification problem, in which a text is to be classified as either positive or negative. However, this task may be approached as a ternary classification problem as well, by introducing a third class of neutral documents. An alternative to such sentiment classification approaches is the determination of a degree of positivity or negativity of natural language text in order to produce, e.g., rankings of positive and negative documents [37,38].

Many state-of-the-art approaches to sentiment classification tasks rely on machine learning techniques [36,39]. On the other hand, some approaches exploit (generic) sentiment lexicons when determining the subjectivity or polarity of natural language text. Both approaches may be combined in hybrid methods as well [29].

In machine learning sentiment analysis methods, natural language text is typically modeled by means of a bag-of-words vector representation, denoting an unordered collection of words occurring in this text. In order to be able to, e.g., distinguish pieces of text from one another in terms of their associated polarity class, machine learning methods typically aim to find and exploit patterns in the vector representations of these texts. In such vector representations, a binary encoding scheme, indicating the presence of specific words, has proven to be effective [39] as well as to outperform frequency-based encoding [40]. Vectors may also contain features other than words, e.g., parts of words, word groups, or features representing semantic distinctions between words [41]. Features represented in vectors may be weighted as well [42].

Lexicon-based methods account for the semantic orientation of individual words in a text by matching these words with a list of words and their associated sentiment scores, i.e., a sentiment lexicon. The text’s overall semantic orientation is then determined by aggregating (e.g., summing) the individual word scores, as retrieved from the sentiment lexicon. Hybrid approaches may realize the aggregation through a machine learning process as well [29]. In this sentiment scoring process, other aspects of content may be taken into account as well, such as...
sentiment detection procedures [26], lexicon-based sentiment analysis methods tend to sacrifice computational efficiency for a classification accuracy which is typically inferior to the classification accuracy of machine learning methods in specific domains for which machine learning approaches can be trained and optimized [30]. However, lexicon-based approaches have an attractive advantage over machine learning methods in that they have a more robust performance across domains and texts [30]. Additionally, lexicon-based approaches enable deep linguistic analysis to be incorporated into the sentiment analysis process [26] which, if fine-tuned, can improve the classification accuracy. Moreover, lexicon-based sentiment scoring approaches are essentially rule-based approaches, which can inherently provide insight into the motivation for the classification of the conveyed sentiment. Last, lexicon-based approaches can be generalized relatively easily to other languages by using dictionaries [35].

2.2. Multi-lingual sentiment analysis

Today’s sentiment analysis systems must deal with an abundance of multi-lingual sentiment-carrying user-generated content. As different approaches are required for distinct languages [44], existing work does not typically focus on devising a single sentiment analysis approach for multiple languages, but rather on analyzing the sentiment conveyed by documents in selected languages, mainly by means of applying sentiment analysis techniques tailored to each specific language. Existing work is primarily focused on how to devise sentiment analysis methods for other languages with minimal effort, without sacrificing too much accuracy. Rather than constructing new frameworks for languages other than the reference language [44–48], recent work focuses on using machine translation techniques in order to be able to re-use many existing tools when performing automated sentiment analysis on multi-lingual natural language content.

Sentiment analysis of machine-translated texts may seem a rather ineffective approach, as machine translation typically fails to correctly translate substantial amounts of text and moreover tends to reduce well-formed texts to sentence fragments. Nevertheless, recent work on sentiment analysis of news messages in nine languages demonstrates that the accuracy of sentiment classification on machine-translated text is largely independent of the quality of the machine translator used (i.e., the translator does not necessarily have to produce well-formed texts) and that sentiment analysis of texts that have been translated into English is rather consistent across languages, after normalizing sentiment scores in order to allow for meaningful cross-cultural comparisons [49].

Other work suggests that in some cases, sentiment analysis of machine-translated texts can yield even better results than sentiment analysis of the original texts [50]. This appears to be the case especially when the original language is not easily interpreted by state-of-the-art natural language processing tools. For instance, in [50], the authors use a Chinese framework for classifying the sentiment of Chinese reviews, and an English framework for classifying the sentiment of Chinese reviews that have been translated into English. The results indicate that sentiment analysis of the translated texts outperforms sentiment analysis of the original texts. An ensemble of both methods further improves the performance.

Machine translation can be utilized in another way as well in order to facilitate automated sentiment analysis in multiple languages. Rather than performing sentiment analysis on machine-translated texts, many researchers focus on automatically generating sentiment lexicons by means of machine translation. A common approach is to automatically translate an existing sentiment lexicon [35], and, possibly, to subsequently propagate the sentiment scores to semantically related words [51]. An alternative approach, which has been shown to outperform machine translation of sentiment lexicons, is to automatically generate a sentiment lexicon from a collection of (automatically) translated and annotated texts [35,52–54]. However, research suggests that the subjectivity of most of the words in sentiment lexicons is lost in translation — subjectivity appears to be a property associated not with words, but with word meanings [35]. Semantic lexicons can be used in order to address this issue.

2.3. Semantic lexicons

A widely used on-line (semantic) lexical resource is WordNet [55], the design of which has been inspired by psycholinguistic theories of human lexical memory. WordNet is organized into sets of cognitive synonyms — synsets — which can be differentiated based on their Part-of-Speech (POS) type. Each WordNet synset expresses a distinct concept and is linked to other synsets through different kinds of relations, such as synonymy, antonymy, hyponymy, or meronymy. The need for such a lexical reference system has arisen as conventional dictionaries do not usually capture such semantic relations. Conventional dictionaries use lexicographical sorting for words for human users’ convenience. Conversely, WordNet has been designed to be used under program control and enables the distinction between different word forms and word meanings.

SentiWordNet [32,33] is a lexical resource in which each WordNet synset is associated with three numerical scores, quantifying its associated sentiment. These scores describe how objective, positive, and negative the terms contained in a synset are. An ensemble of eight ternary classifiers has been used to classify each synset as either objective, positive, or negative, based on a vector representation of the associated description of the synset. The overall objectivity, positivity, and negativity scores for a synset have then been determined by the (normalized) proportion of classifiers that assigned the corresponding labels to the synset.

The availability of semantic lexical resources is not limited to the English language. For instance, EuroWordNet [56] has been developed as a collection of semantic lexicons for several European languages, including English, Dutch, Italian, and Spanish. For each supported language, a semantic lexicon has been created, with a structure similar to the structure of WordNet. Additionally, EuroWordNet has been designed in such a way that the language-specific semantic lexicons are linked to one another through WordNet, such that each English synset is associated with its equivalents in the languages included in EuroWordNet.

For Dutch, i.e., the language considered in our current work as an alternative to our English reference language, a more extensive semantic lexicon has been developed on top of EuroWordNet as well. In DutchWordNet (Cornetto) [57], the Dutch part of EuroWordNet has been enriched with information from the Referentie Bestand Nederland (RBN), which is a lexical database for Dutch, containing information on orthography, morphology, syntax, semantics, pragmatics, and combinatorics.

Language-specific semantic lexical resources and their interlinkage through semantic lexical resources such as WordNet can facilitate new approaches for extending an existing lexicon-based sentiment analysis approach from one language to another. The semantic relations between language-specific semantic lexicons could be exploited in order to propagate a sentiment lexicon from one language to another, while preserving semantics. Alternatively, sentiment scores for a seed set of words could be propagated through a language-specific semantic lexicon in order to generate language-specific sentiment lexicons [34,58–60]. As both types of approaches account for semantics, they may compensate for the drawbacks of existing machine translation methods for multi-lingual sentiment analysis.

3. Framework

In order to investigate how lexicon-based sentiment analysis can be extended from our reference language, i.e., English, to another language, i.e., Dutch, we first need a lexicon-based sentiment analysis framework
for the reference language. This framework can then serve as a starting point for an extension to another language.

3.1. Polarity classification

Building upon our previous work [34], we use a binary polarity classifier that classifies documents as either positive or negative based on the aggregated sentiment scores for individual words, as retrieved from a semantic sentiment lexicon such as SentiWordNet. For an arbitrary synset, we compute a single sentiment score based on its objectivity, positivity, and negativity scores (all positive real numbers which sum to 1), by subtracting the negativity score from the positivity score, thus obtaining a real number in the interval $[-1, 1]$, representing sentiment scores in the range from negative to positive, respectively.

In our polarity classification process, detailed in Algorithm 1, documents are first split into sentences and words. Then, each word’s POS tag, lemma, and word sense are determined in order to subsequently retrieve its sentiment score from the sentiment lexicon. For the word sense disambiguation process, we use a Lesk-based algorithm for WordNet [61], as described in our previous work [34]. For each word in a sentence, the algorithm essentially selects the word sense that is semantically most similar to the words in the context, i.e., the other words in a sentence.

After retrieving all word-level sentiment scores from the sentiment lexicon, the sentiment score $s_\delta$ of a document $d$ is computed by summing the sentiment scores $s_i$ of each non-stopword $t$ in each sentence $s$ of the document, i.e.,

$$s_\delta = \sum_{s \in d} \sum_{t \in s} s_i,$$

(1)

The resulting document-level sentiment score is subsequently used in order to classify the document’s polarity class $c_d$ as either positive (1) or negative ($-1$), i.e.,

$$c_d = \begin{cases} 1 & \text{if } (s_\delta - \epsilon) \geq 0, \\ -1 & \text{if } (s_\delta - \epsilon) < 0, \end{cases}$$

(2)

with $\epsilon$ representing an offset that corrects for a possible bias towards positivity in sentiment scores. Such a bias may be caused by people’s tendency to write negative texts with rather positive words [30]. Following existing work [30], we calculate the offset $\epsilon$ on a training set as

$$\epsilon = 0.5 \left( \frac{\sum_{\phi \in \Phi} s_{\phi,d} \delta_{\phi,d}}{|\Phi|} + \frac{\sum_{\phi \in \Phi} s_{\phi,d} (1 - \delta_{\phi,d})}{|\Phi|} \right),$$

(3)

with $\Phi$ denoting the subset of positive documents in the training set, and $\Phi$ denoting the subset of negative documents in the training set.

Algorithm 1. Classifying a document’s polarity

| input : A document $d$ and an offset $\epsilon$ |
| output: The polarity classification $c_d$ of document $d$ |
| 1 $s_\delta = 0$; |
| 2 foreach $s \in d$ do |
| 3 foreach $t \in s$ do |
| 4 $pos = findPOS(t, s);$ |
| 5 $lemma = findLemmas(t, pos);$ |
| 6 $sense = findWordSense(t, s, pos);$ |
| 7 $s_i = getWordScore(lemma, sense, pos);$ |
| 8 $s_\delta = s_\delta \delta + s_i;$ |
| 9 end |
| 10 end |
| 11 $c_d = 1;$ |
| 12 if $(s_\delta - \epsilon) < 0$ then |
| 13 $c_d = -1;$ |
| 14 end |
| 15 return $c_d$; |

Our sentiment analysis framework has been developed for classifying the polarity of English documents. As such, in order to be able to classify the polarity of documents written in another language, the latter documents could be automatically translated into the reference language, such that they can be analyzed by means of the sentiment analysis framework for the reference language. Thus, our existing English sentiment analysis framework can be used for classifying the polarity of Dutch documents without needing to develop any new natural language processing components other than a machine-translation component.

However, the concepts of our framework can be used for polarity classification in Dutch as well, if lexical and syntactical parsing tools for identifying sentences, words, POS, and lemmas are available for Dutch, as well as a semantic lexical resource for the Dutch language. The latter semantic lexical resource can be used for word sense disambiguation, as well as for constructing a Dutch sentiment lexicon that can be used in a sentiment analysis framework with components tailored to the Dutch language.

Our framework (visualized in Fig. 1) supports two of such alternatives to the machine translation approach. First, we consider traversing the relations between language-specific semantic lexicons in order to map the existing sentiment lexicon for the English reference language to a new sentiment lexicon for the Dutch target language. This method is detailed in Section 3.2. Second, we consider propagating sentiment within language-specific semantic lexical resources, as described in Section 3.3.

3.2. Traversing relations between language-specific semantic lexical resources

The valuable information contained in the sentiment lexicon of an existing sentiment analysis approach for the reference language can be utilized in another language when it is used to generate a sentiment lexicon for the target language. This may be done by (automatically) translating an existing sentiment lexicon from the reference language into the target language [35,51]. However, as subjectivity tends to be associated with word meanings rather than words [35], we propose a novel method of translating a sentiment lexicon from a reference language to a target language, while taking into account the semantics of the words in the sentiment lexicons. To this end, we exploit language-specific semantic lexical resources and their interrelations.

In our novel cross-lingual sentiment score mapping method SMAP (see Fig. 2), we assume an existing sentiment lexicon for the reference language to be linked to a semantic lexical resource with meaningfully related words and concepts (synsets). Provided that a mapping exists between this semantic lexicon and an equivalent semantic lexicon for another language, the sentiment from the reference sentiment lexicon can be mapped to a new sentiment lexicon for the target language by traversing the associated relations between the semantic lexicons of both respective languages.

For example, for our reference language (English) and target language (Dutch), the English SentiWordNet sentiment lexicon can be used as a starting point for our proposed cross-lingual sentiment score mapping procedure. SentiWordNet contains sentiment scores for all synsets in the WordNet semantic lexicon. Additionally, a mapping exists between WordNet and its Dutch equivalent DutchWordNet (Cornetto). By exploiting these relations, SentiWordNet sentiment scores associated with English WordNet synsets can be projected onto equivalent Dutch synsets in DutchWordNet (Cornetto), thus yielding a Dutch sentiment lexicon.

In order to propagate sentiment scores associated with synsets through language-specific semantic lexical resources, we first map the reference sentiment lexicon’s synsets to the reference semantic lexicon. Subsequently, we map the synsets in the semantic lexicon for the new language to their equivalent synsets in the reference semantic lexicon. Then, for each synset in the reference sentiment lexicon, we use these mappings to assign the associated reference sentiment score to the equivalent synsets and their synonyms in the semantic lexicon for the
target language. The result is saved in the new sentiment lexicon for the target language. This process is further detailed in Algorithm 2.

Algorithm 2. Sentiment propagation through relations between semantic lexical resources

```plaintext
input : The reference sentiment lexicon $S^*$, the reference semantic lexicon $L^*$, and the semantic lexicon for the target language $L'$
output: The sentiment lexicon for the target language $S'$
1. $S' = \emptyset$
2. foreach sentiWord $\in S^*$ do
3.   synset* = getSynset(sentiWord, $S^*$);
4.   synset' = mapSynsetFromTo(synset*, $L^*$, $L'$);
5.   if synset' $\neq \emptyset$ then
6.     pos = getPos(synset*, $L^*$);
7.     $\zeta$ = getScore(synset*, $S^*$);
8.     synonyms = getSynonyms(synset*, $L^*$);
9.     foreach $t \in$ synonyms do
10.    lemma = getLemma($t$, $L^*$);
11.    sense = getWordSense($t$, $L^*$);
12.    $S'$ = ($S'$, (synset', lemma, sense, pos, $\zeta$));
13. end
14 end
15 return $S'$;
```

3.3. Sentiment propagation within language-specific semantic lexical resources

When creating a new sentiment lexicon for a new target language, one could also consider not using the reference sentiment lexicon as a starting point, as the sentiment associated with words or word meanings may have a cultural dimension. Instead, one could consider creating a new sentiment lexicon for the target language by means of an approach involving propagating the sentiment of a small seed set of words to words which are semantically related [34,58–60].

In our sentiment propagation method SPROP (detailed in Algorithms 3 and 4 and visualized in Fig. 3), semantic relations in a language-specific semantic lexicon are traversed for each seed word. Examples of such semantic relations are hyponymy (type-of relations), synonymy, and antonymy. In the sentiment propagation process, each encountered word $t$ is stored with a sentiment score $\zeta_t$, based on the score $\xi_t$ of the seed word, a diminishing factor $\delta$, and the number of steps $k$ (with a maximum of $K$) between the seed word and $t$, i.e.,

$$\zeta_t = \xi_t \delta^k, \quad \tau \epsilon \{-1, 1\}, \quad k \epsilon \{1, \ldots, K\}, \quad -1 \leq \xi \leq 1, \quad 0 < \delta \leq 1, \quad (4)$$

with $\tau$ indicating whether to invert $(-1)$ the score, i.e., when traversing antonym relations, or not $(1)$.

In each iteration of our algorithm, the computed sentiment score $\zeta_t$ for the current word $t$ is propagated to the words in its directly related synsets. While doing so, with each next traversed semantic relation, the propagated sentiment is further diminished. As a result, words that are semantically more closely related to a seed word obtain a higher absolute sentiment score than those with a more indirect semantic relation to a seed word. If a word is encountered multiple times when propagating the sentiment associated with seed words, this word is assigned the score obtained from the shortest path between the word and any of the seeds, because we
assume that the shorter the path, the more accurate the sentiment can be determined.

4. Evaluation

An evaluation of the performance of the methods proposed in Section 3 can provide insight in how lexicon-based sentiment analysis can best be extended from our reference language, i.e., English, to another language, i.e., Dutch. This evaluation can help understand the importance of semantic relations – both across and within languages – for multi-lingual sentiment analysis. In this section, we present our experimental set-up and discuss our experimental results.

4.1. Experimental setup

We focus on 600 positive and 600 negative opinionated Dutch documents on 40 distinct topics, crawled from Dutch review Web sites, forums, and blogs. The documents have been classified by three human annotators, until they reached full consensus. On this corpus, we assess the performance of our considered methods by means of the 10-fold cross-validated overall sentiment classification accuracy and the macro-level $F_1$-score. We assess the statistical significance of performance differences by means of a paired two-sample one-tailed t-test.

The implementation of our sentiment classification framework has been done in C#.Net. We have built upon our existing framework for classifying the sentiment of English documents [34], which classifies sentiment as described in Section 3.1. We have constructed a similar implementation for sentiment classification of Dutch documents, which is an extension of the English implementation by means of the translation and sentiment propagation methods discussed in Section 3.

For classifying the sentiment of English text, our implementation uses regular expressions in order to split the text into words. POS tagging is done with a SharpNLP [62] POS tagger. Lemmatization

![Fig. 2. Our novel method for mapping sentiment scores from a reference language to a target language. Positive words and synsets are marked with vertical stripes, whereas negative words and synsets are marked with horizontal stripes. Others are left blank. Darker shading implies stronger sentiment.](image)

![Fig. 3. Our proposed method for propagating the sentiment of a set of seed words through the semantic lexicon of a target language. The sentiment lexicon thus generated in one propagation step is visualized. Positive words and synsets are marked with vertical stripes, whereas negative words and synsets are marked with horizontal stripes. Others are left blank. Darker shading implies stronger sentiment.](image)
and word sense disambiguation is performed by means of the C# WordNet.Net [63] WordNet API. The sentiment classification process relies on a semantic lexicon and a sentiment lexicon. We link English word senses to WordNet [55], whereas we retrieve the associated sentiment scores from SentiWordNet 3.0 [33]. On a widely used data set of 1000 positive and 1000 negative English movie reviews [39], our implementation has an overall sentiment classification accuracy and macro-level F1-score of approximately 60% [34].

**Algorithm 3.** Propagating sentiment in a language-specific semantic lexical resource

```plaintext```
input : The target semantic lexicon L', a list seeds of the words to propagate, their associated scores \( s \), the maximum number of iterations \( K \), and a diminishing factor \( \delta \)
output : A sentiment lexicon \( S' \) containing all propagated words with their computed sentiment scores
1. \( s^\prime(\text{syn}) = \text{getSynsets}(L') \)
2. \( S' = \emptyset \)
3. foreach \( t \in \text{seeds} \) do
4. \( s = \text{score}(t) \)
5. \( S' = \text{propWord}(s, S', t, s, \delta, K) \) // Alg. 4
6. end
7. return \( S' \)
```

**Algorithm 4.** Propagating a single word’s sentiment in a lexical resource (propWord)

```plaintext```
input : All synsets in the semantic lexicon for the target language, a sentiment lexicon for the target language \( S' \), a word \( t \) to propagate, the score \( s = \text{score}(t) \), a diminishing factor \( \delta \), the current iteration \( k \), and the maximum number of iterations \( K \)
output : A sentiment lexicon \( S' \) containing all propagated words with their computed sentiment scores
1. if \( k < K \) then
2. \( \text{reached} = \text{getSteps}(t); // \text{do for new} \ t \)
3. if \( \text{reached} > k \) then
4. \( \text{synsetsWithWord} = \text{getSynsets}(<\text{synset}, \text{syn}>, t); \)
5. foreach \( \text{synset} \in \text{synsetsWithWord} \) do
6. \( \text{pos} = \text{getPOS}(\text{synset}); \)
7. \( \text{syns} = \text{getSynonyms}(\text{synset}); \)
8. foreach \( \text{syn} \in \text{syns} \) do
9. \( \text{lemmas} = \text{getLemmas}(\text{syn}); \)
10. \( \text{sense} = \text{getWordSense}(\text{syn}); \)
11. \( \text{scores} = \{\text{score}, \text{sense}, \text{pos}, s\} \)
12. end
13. \( \text{rels} = \text{getRelations}(\text{synset}); \)
14. \( \text{foreword} \in \text{rels} \) do
15. \( \text{tau} = 1 \)
16. if \( r = \text{antonym} \) then
17. \( \text{tau} = -1 \)
18. end
19. \( \text{rSyns} = \text{getSynonyms}(r); \)
20. \( \text{foresword} \in \text{rSyns} \) do
21. \( \text{tau} = s \)
22. \( k' = k + 1 \)
23. \( \text{theta} = \{\text{synsets}, \text{syns}, \text{tau}, s, \text{pos}, \text{score}\} \)
24. \( \text{S'} = \text{propWord}(\text{theta}); \)
25. end
26. end
27. end
28. \( \text{setSteps}(t, k); \)
29. end
30. return \( S' \)
```

The implementation of our sentiment classification method for Dutch text is similar to the implementation for English text, even though it utilizes different language-specific components. For POS tagging in Dutch, we use a SharpNLP [62] POS tagger. Lemmatization is performed by the Tadpole [64] lemmatizer. Word sense disambiguation is done by applying our own implementation of the Lesk-based algorithm implemented in WordNet.Net [63]. The Dutch sentiment classifier relies on DutchWordNet (Cornetto) [57], a large semantic lexical resource for Dutch, which is used for word sense disambiguation as well as for sentiment lexicon creation using one of our considered methods other than our machine translation baseline.

### Table 1

Seed sets of sentiment-carrying words.

<table>
<thead>
<tr>
<th>Seed word</th>
<th>Score</th>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mooi</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Schoon</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Anbiddelijk</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Duidelijk</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Elegant</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Beter</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Glimlend</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Perfect</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Energiëk</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Trots</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Super</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Scherterend</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hart</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Amicaal</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Gezelheid</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Goeid</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Anbidden</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Plezier</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Aangenaam</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Uiltuend</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Beeldig</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Positief</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Veilig</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Vrijheid</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Vakantie</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Ootspanning</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Kote</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Boos</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Arrogant</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Bewolkt</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Vreemd</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Onmogelijk</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Haat</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Twijfelen</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Verraad</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Imbeciel</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mongool</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Tering</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Wantrouwelig</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Vervold</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Gedachteleenos</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Berucht</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Jammer</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Treurig</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Onheilspellend</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Onheilspellend</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Schelden</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Irriteren</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Vervelen</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Negatief</td>
<td>1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### Table 2

Performance of our approaches, based on a 10-fold cross-validation on our data set. The best performance is printed in bold for each performance measure.

<table>
<thead>
<tr>
<th>Method</th>
<th>Positive Precision</th>
<th>Positive Recall</th>
<th>Positive F1</th>
<th>Negative Precision</th>
<th>Negative Recall</th>
<th>Negative F1</th>
<th>Overall Precision</th>
<th>Overall Recall</th>
<th>Overall F1</th>
<th>Accuracy</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT</td>
<td>0.416</td>
<td>0.385</td>
<td>0.400</td>
<td>0.428</td>
<td>0.460</td>
<td>0.443</td>
<td>0.423</td>
<td>0.422</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMap</td>
<td>0.547</td>
<td>0.500</td>
<td>0.523</td>
<td>0.540</td>
<td>0.587</td>
<td>0.562</td>
<td>0.543</td>
<td>0.542</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SProp 1</td>
<td>0.428</td>
<td>0.397</td>
<td>0.412</td>
<td>0.438</td>
<td>0.470</td>
<td>0.453</td>
<td>0.433</td>
<td>0.433</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SProp 2</td>
<td>0.596</td>
<td>0.582</td>
<td>0.591</td>
<td>0.605</td>
<td>0.598</td>
<td>0.593</td>
<td>0.593</td>
<td>0.593</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SProp 3</td>
<td>0.633</td>
<td>0.578</td>
<td>0.605</td>
<td>0.612</td>
<td>0.665</td>
<td>0.637</td>
<td>0.622</td>
<td>0.621</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We consider three main sentiment analysis approaches. In our machine translation (MT) baseline, first, we automatically translate the Dutch texts from our considered corpus into English by using the Google Translate service [65]. Then, we classify the sentiment conveyed by the translated documents by means of our sentiment classification approach for English documents.

Our first alternative to this machine translation baseline is a cross-lingual sentiment score mapping method (SMAP), in which we first map the sentiment associated with all WordNet synsets from SentiWordNet 3.0 to all equivalent synsets in DutchWordNet (Cornetto). We subsequently classify the sentiment conveyed by the Dutch documents in our corpus by means of our sentiment classification approach for Dutch text, while utilizing the Dutch sentiment lexicon thus constructed.

As a second alternative to the machine translation baseline approach, we use the SPROP method in order to propagate the sentiment of a set of seed words through DutchWordNet (Cornetto) and subsequently classify the conveyed sentiment by using the constructed sentiment lexicon in our sentiment classification method for Dutch documents. We assess SPROP with three distinct seed sets, containing positive words (with a sentiment score of 1) and negative words (with a sentiment score of -1). For each of these seed sets, sentiment scores are propagated by traversing the holonym, hyperonym, and hyponym relations between synsets in DutchWordNet (Cornetto), with a maximum number of iterations $K$ of 8 and a diminishing factor $\delta$ of 0.9, as an initial optimization of these parameters by means of a hill-climbing procedure on one fold of our data indicated that these settings were most promising.

Each of our seed sets, detailed in Table 1, has been manually constructed by our three human annotators, all of whom are native Dutch speakers. The human annotators have combined their knowledge of the Dutch language with the most positive and negative synsets in SentiWordNet in order to construct seed sets for a Dutch sentiment lexicon. The first set contains ten positive and ten negative Dutch words. The second set is an expansion of the first set, such that it contains 26 positive and 17 negative Dutch words. Another, final expansion has resulted in a third seed set, containing 26 positive and 24 negative Dutch words.

| Table 3 | Relative differences of the 10-fold cross-validated overall accuracy of our approaches, benchmarked against one another on our collection of Dutch documents. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Benchmark       | MT              | SMAP            | SPROP 1         | SPROP 2         | SPROP 3         |
| MT              | 0.000           | 0.286***        | 0.026           | 0.404***        | 0.471***        |
| SMAP            | -0.223***       | 0.000           | -0.202***       | 0.092***        | 0.144**         |
| SPROP 1         | -0.025          | 0.254***        | 0.000           | 0.369***        | 0.415**         |
| SPROP 2         | -0.288***       | -0.084**        | -0.270***       | 0.000           | 0.048**         |
| SPROP 3         | -0.320***       | -0.136***       | -0.303***       | -0.046**        | 0.000           |

Performance differences marked with * are statistically significant at $p < 0.05$, those marked with ** are significant at $p < 0.01$, and those marked with *** are significant at $p < 0.001$.

4.2. Experimental results

The performance of our methods of classifying the sentiment conveyed by Dutch documents by exploiting an existing method for sentiment classification of English documents is summarized in Tables 2, 3, and 4. These experimental results demonstrate that some of our approaches work better than others for performing sentiment analysis of documents in another language than the reference language. Several observations can be made in Tables 2, 3, and 4.

In general, all approaches exhibit a rather balanced performance, as they seem to perform equally well when classifying the sentiment of positive and negative documents. Additionally, when exploiting our existing sentiment analysis framework for English texts by means of our considered approaches, the best achievable performance of our framework on Dutch documents is rather comparable to the performance of the existing framework on English documents.

As reported in our previous work [34], the existing English sentiment analysis approach can obtain an overall accuracy and macro-level $F_1$-score of about 60% on a widely used collection of English movie reviews [39]. The machine translation (MT) baseline yields a sentiment classification performance on Dutch documents that is inferior to the reported performance on English documents. When using the MT method, we obtain an overall accuracy and macro-level $F_1$-score around a mere 47%. The SMAP method yields an overall sentiment classification accuracy and macro-level $F_1$-score of about 54% on Dutch documents, whereas these scores amount to about 62% for SPROP.

The experimental results on our corpus of Dutch documents show that our novel cross-lingual sentiment score mapping method (SMAP) significantly outperforms our machine translation (MT) baseline by about 29%, caused by increased precision and recall for both positive and negative documents. Clearly, valuable information on sentiment is (partially) contained in the semantics of our source language (i.e., English), and is as such preserved when accounting for semantics by mapping the sentiment lexicon to our target language (i.e., Dutch) through relations between language-specific semantic lexicons. Accounting for semantics when propagating the sentiment of a seed set of sentiment-carrying...
words within a language (SPROP) has even greater potential than exploiting semantics when mapping sentiment across languages. SPROP significantly outperforms both MT and SMAP by up to about 47% and 14%, respectively. This suggests that sentiment is not only linked to word meanings, but tends to be language-specific as well.

The machine translation approach may be thwarted by text meaning getting lost in translation. With the SMAP method, noise may be introduced on word-level meanings, which apparently do not only depend on semantics, but can also be language-specific. The SPROP method is insensitive to such translation errors, as it depends on language-specific seed sets of sentiment-carrying words. The advantage of SPROP does however appear to depend on the set of seed words used in the lexicon creation process. Our results suggest a sensitivity of the sentiment classification performance to the size of the seed set.

The smallest seed set, i.e., seed set 1, does not yield significant improvements over any of our methods. Conversely, a somewhat larger seed set, i.e., set 2, yields significant improvements over the MT baseline and the SPROP 1 method, as well as a small, yet significant improvement over SMAP. Set 3, i.e., the largest seed set, yields the largest, significant improvements over MT, SMAP, SPROP 1, and SPROP 2. This may be explained by a larger part of the sentiment lexicon being manually annotated (i.e., the sentiment-carrying words in the seed sets), as well as by such larger initial lexicons being expanded to larger sentiment lexicons.

That the SPROP 2 and SPROP 3 sentiment lexicons are comparably large is clearly visible in Fig. 4, where SPROP 1, SPROP 2, and SPROP 3 respectively cover 13%, 24%, and 29% of all terms in DutchWordNet (Cornetto), while covering 8%, 18%, and 20%, respectively, of all terms occurring in the corpus. Interestingly, the SMAP lexicon yields a significantly better performance than the SPROP 1 lexicon, even though the SMAP lexicon only covers about 8% of the words in the corpus as well (albeit a different subset). Moreover, while covering more than two times as many of the terms occurring in the corpus, the SPROP 2 lexicon significantly outperforms the SMAP lexicon with only about 9%. Hence, the sentiment-carrying words in the SPROP lexicon, constructed by exploiting semantic relations between languages, are comparably valuable in the analysis of the sentiment conveyed by our Dutch documents. This suggests that not only the size, but also the suitability of the seed sets for the corpus matters.

Fig. 4 additionally shows that the SPROP lexicons mostly cover a different part of the terms in DutchWordNet (Cornetto) than the SMAP lexicon. Especially the larger SPROP lexicons cover a large part of the space, in addition to the 24%, 35%, and 40% coverage of the SMAP lexicon by the respective SPROP 1, SPROP 2, and SPROP 3 lexicons. The extra coverage of the larger SPROP lexicons helps improve their performance over the SMAP lexicon. This confirms the importance of exploiting semantic relations within a language when constructing a sentiment lexicon.

A failure analysis has revealed that the SPROP approach occasionally fails where SMAP succeeds. This tends to happen when analyzing the sentiment conveyed by texts containing sentiment-carrying words that have not been assigned appropriate scores in the sentiment score propagation process. SPROP may have failed to assign an appropriate sentiment score because either the associated synset was not reached by the propagation process, or the sentiment score was significantly diminished because the distance of the synset to the (possibly non-optimal) seed words was too large. Additionally, we have encountered cases in which the SMAP method fails, where the SPROP variants succeed. This happens when the SMAP mappings do not capture the true semantics of words in Dutch, whereas the propagated SPROP lexicons approximate this better.

Our failure analysis has additionally revealed that, occasionally, all of our methods fail because of misinterpreting texts. Such misinterpretations typically occur in the case of negation or amplification of sentiment. Additionally, sarcasm and proverbs are interpreted literally by our current methods, as they are not covered by the resources available. Hashtags and other (misspelled) terms that are neither in our semantic lexicon nor in the constructed sentiment lexicons are another source of misinterpretations. Last, more complex structures of sentences, paragraphs, and documents are not currently taken into account. As these structures constitute the way in which sentiment-carrying words convey an author’s sentiment, not accounting for these structures can cause a misinterpretation of the text in terms of its conveyed sentiment.

5. Conclusions

We have explored several methods of expanding an existing lexicon-based sentiment analysis method for a reference language, i.e., English, to another language, i.e., Dutch. Our findings suggest that, when analyzing the sentiment conveyed by texts in the target language, we cannot rely on an existing, well-performing sentiment lexicon for the reference language when simply machine- translating texts to the reference language and subsequently using the existing method for classifying the sentiment of the translated texts.

Conversely, when we map sentiment from the well-performing sentiment lexicon for our reference language to the target language by exploiting relations between language-specific semantic lexicons, we can achieve significantly better sentiment classification performance in the target language. Accounting for semantics by propagating the sentiment of a seed set of sentiment-carrying words to semantically related words within the target language has even greater potential, provided that the seed set of sentiment-carrying words is sufficiently large. This indicates that sentiment is not only linked to word meanings, but also tends to have a language-specific dimension. Thus, semantics could be exploited within a language, in addition to their use as a universal link between languages when constructing sentiment lexicons in a target language.

Nevertheless, our novel sentiment mapping method, exploiting relations between language-specific semantic lexicons, has two attractive advantages over the alternative sentiment propagation method. First, in order for sentiment propagation to be truly effective, a large set of seed words in the target language is needed, whereas our sentiment mapping method does not need a seed set at all. Second, sentiment propagation is computationally more complex than our sentiment mapping method.

All in all, the key insight brought forward by our work is that semantic relations between and within languages should be carefully considered in order to exploit the full potential of sentiment analysis in real-life decision support systems that support natural language content in multiple languages. With the accuracy levels that can be obtained by our semantics-guided methods, sentiment-related information that is extracted from text in other languages than the reference language can be presented to decision makers as a rough indication of where their attention may be needed.

Our findings warrant several directions for future work. First, we could validate our findings on data in another target language. Another possible direction for future research would be to further optimize the seed sets used for the sentiment propagation process, such that they, e.g., maximize the coverage of the exploited semantic lexicon. Additionally, in future work, we could explore how to combine the sentiment propagation process with our proposed semantics-guided cross-lingual sentiment mapping approach in order to best exploit the strengths of both approaches. Last, as our findings indicate that sentiment tends to be partly language-specific, we aim to explore the comparability of sentiment scores across languages, as well as how such language-specific sentiment scores relate to an author’s intended sentiment.

Acknowledgments

The authors of this paper are partially supported by the Dutch national program COMMIT. Special thanks go to Teezir (http://www.teezir.com) for their data and technical support.


A. Hogenboom holds a B.Sc. degree and a cum laude M.Sc. degree in economics and informatics, obtained at Erasmus University Rotterdam, The Netherlands, in 2007 and 2009, respectively. Alexander is currently a Ph.D. student at the Erasmus University Rotterdam. He is affiliated with the Econometric Institute, the Research Center for Business Intelligence at the Erasmus Research Institute of Management, and Erasmus Studio. In his research, Alexander explores the utilization of methods and techniques from informatics for facilitating and supporting decision making processes. His research interests relate to intelligent systems for information extraction, focused on tracking and monitoring of economic sentiment.

Flavius Frasincar obtained an M.Sc. degree in computer science from Politehnica University Bucharest, Romania, in 1998. In 2000, he received a professional doctorate degree in software engineering from the Eindhoven University of Technology, The Netherlands. He got a Ph.D. degree in computer science from the Eindhoven University of Technology, The Netherlands, in 2005. Since 2005, he is an assistant professor in information systems at the Econometric Institute of Erasmus University Rotterdam, The Netherlands. He has published in numerous conferences and journals in the areas of databases, Web information systems, personalization, and the Semantic Web. He is a member of the editorial board of the International Journal of Web Engineering and Technology.

Bas Heerschop obtained both a B.Sc. degree and an M.Sc. degree in economics and informatics at the Erasmus University Rotterdam, The Netherlands, in 2010 and 2012, respectively. His research interests cover various aspects of sentiment analysis, ranging from accounting for negation and structural aspects of text when analyzing its conveyed sentiment to sentiment lexicon creation and sentiment analysis in a multi-lingual setting.

Uzay Kaymak received an M.Sc. degree in electrical engineering, a degree as a chartered designer in information technology, and a Ph.D. degree in control engineering from the Delft University of Technology, Delft, The Netherlands, in 1992, 1995, and 1998, respectively. From 1997 to 2000, he was a reservoir engineer with Shell International Exploration and Production. Currently, he is a professor of information systems in health care at the Department of Industrial Engineering & Innovation Sciences of the Eindhoven University of Technology. The Netherlands. Prof. Kaymak is an associate editor of IEEE Transactions on Fuzzy Systems and is a member of the editorial board of several journals.

Franciska de Jong is a professor of language technology at the University of Twente, The Netherlands, since 1992. She is also affiliated with the Erasmus University Rotterdam, The Netherlands, where she is the director of the Erasmus Studio. She studied Dutch language and literature at the University of Utrecht, The Netherlands, did a Ph.D. track in theoretical linguistics and started to work on language technology in 1985 at Philips Research, where she worked on machine translation. Currently, her main research interest is in the field of multimedia indexing, text mining, semantic access, cross-language retrieval, and the disclosure of cultural heritage collections (in particular spoken audio archives). She is frequently involved in international program committees, expert groups, and review panels, and has initiated a number of European Union projects. Since 2008, she is a member of the Governing Board of The Netherlands Organization for Scientific Research (NWO).