A Linked Open Data Approach for Sentiment Lexicon Adaptation

Hassan Saif,[1] Miriam Fernandez,[1] Leon Kastler,[2] and Harith Alani[1]

[1] Knowledge Media Institute, The Open University, United Kingdom
{h.saif, m.fernandez, h.alani}@open.ac.uk
[2] Institute for Web Science and Technology, University of Koblenz-Landau, 56070 Koblenz, Germany
lkastler@uni-koblenz.de

Abstract. Social media platforms have recently become a gold mine for organisations to monitor their reputation by extracting and analysing the sentiment of the posts generated about them, their markets, and competitors. Among the approaches to analyse sentiment from social media, approaches based on sentiment lexicons (sets of words with associated sentiment scores) have gained popularity since they do not rely on training data, as opposed to Machine Learning approaches. However, sentiment lexicons consider a static sentiment score for each word without taking into consideration the different contexts in which the word is used (e.g, *great* problem vs. *great* smile). Additionally, new words constantly emerge from dynamic and rapidly changing social media environments that may not be covered by the lexicons. In this paper we propose a *lexicon adaptation* approach that makes use of semantic relations extracted from DBpedia to better understand the various contextual scenarios in which words are used. We evaluate our approach on three different Twitter datasets and show that using semantic information to adapt the lexicon improves sentiment computation by 3.7% in average accuracy, and by 2.6% in average F1 measure.

Keywords: Sentiment Lexicon, Linked Open Data, Twitter

1 Introduction

Sentiment analysis is nowadays an integral part of social listening, used by companies, individuals, governments and organisations to track the sentiment and opinions expressed in social media platforms such as Twitter or Facebook. Most existing approaches to sentiment analysis rely on general-purpose sentiment lexicons (sets of words with associated sentiment scores) to compute the sentiment of a text regardless of its context or domain [3, 13, 20, 7]. However, a word's sentiment may vary according to the context in which the word is used [22]. For example, the word *great* conveys different sentiment when associated with the word *problem* than with the word *smile*. Therefore, the performance of these lexicons may drop when used to analyse sentiment over specific domains.

Some works have attempted to address this problem by adapting pre-built sentiment lexicons to fit new specific domains or contexts [6, 19, 8, 18]. In our previous work [18] we proposed a lexicon adaptation approach that uses the *contextual semantics* of words (i.e., semantics inferred from the co-occurrence patterns of words [24]) to capture their context in text and update their prior sentiment orientations and/or strengths in the given sentiment lexicon accordingly. Lexicon adaptation based on words' context has proved to: (i) improve the coverage and applicability of these lexicons on data from different domains, and (ii) enhance sentiment analysis performance in comparison with using sentiment lexicons without adaptation.

While useful, *contextual semantics* may be insufficient, particularly when extracted from short, noisy and ill-structured sentences that frequently appear in social media data. In this work, we build on our previous lexicon adaptation approach [18] and enrich it by exploiting Linked Data, and particularly relations between entities and concepts in DBpedia, to better capture the context of words and update their sentiment orientation and/or strength in the sentiment lexicon accordingly.

Our hypothesis here is that *conceptual semantics*, i.e., semantics extracted from knowledge graphs such as DBpedia, can help to better capture the domain or context for which the lexicon is being adapted, thus aiming to contribute towards a more informed calculation of words' sentiment weights. For example, the context of the word "`Ebola`" in "`Ebola continues spreading in Africa!`" does not indicate a clear sentiment for the word. However, "`Ebola`" is associated with the semantic type (concept) "`Virus/Disease`", which suggests that its sentiment is likely to be negative.

We evaluate our lexicon adaptation model over three Twitter datasets, and show an average, statistically significant, improvement of 3.7% in accuracy, and 2.6% in F1, against the baseline methods when adapting Thelwall-Lexicon, the state-of-the-art sentiment lexicon on social media.

The remainder of this paper is structured as follows. Related work is discussed in Section 2. Our general model for sentiment lexicon adaptation and its semantic enrichment is presented in Section 3. Experimental setup and results are presented in Sections 4 and 5 respectively. Conclusions are reported in Section 6.

2 Related Work

General purpose sentiment lexicons (e.g., MPQA[26], SentiWordNet[2], Thelwall-lexicon[21]) have been traditionally used in the literature to determine the overall sentiment of texts. These lexicons capture a selection of popular words and their associated weighted sentiment orientations, without considering the domain, topic, or context where the lexicons are being used. However, a word's associated sentiment may vary according to the context in which the word is used [22]. To address this problem multiple works have emerged in recent years to: (i) create domain-specific lexicons or, (ii) adapt existing lexicons to specific domains.

Most of the existing works belong to the first category, where approaches have been proposed to develop sentiment lexicons tailored for specific domains [11, 19, 9]. However, several approaches have proposed methods for adapting existing, well-known lexicons, to specific domains [6, 15, 18]. As previously mentioned, lexicon adaptation not only reduces the burden of creating lexicons from scratch, but also supplements the process with a collection of pre-existing words and their sentiment orientations and weights. While the majority of work on lexicon adaptation focuses on conventional text, lexicon adaptation for social media data is still in its infancy. One very recent work in this line [8] has focused on updating the sentiment of neutral words in SentiWordNet. In addition to this work, we not only adapt sentiment weights, but also study the extraction and addition of new terms not provided in the original lexicon [18]. This is potentially useful in the case of social media data, where new terms and abbreviations constantly emerge. Note that, in-line with the work of Lu and colleagues [11], our proposed lexicon adaptation method is not restricted to domain-adaptation, but rather considers a more fine-grained context adaptation, where the context is defined by a collection of posts. Moreover, our approach does not make use of training data to adapt the lexicon.

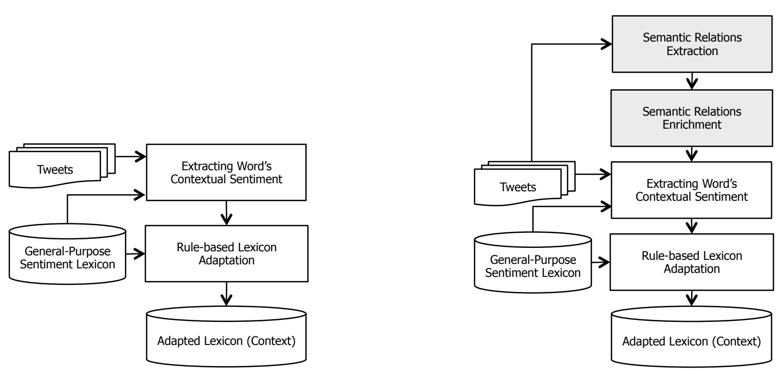
Another novelty of our approach with respect to previous works, is the use of conceptual semantics, i.e. semantics extracted from ontologies such as DBpedia, to adapt sentiment lexicons. Our hypothesis is that conceptual semantics can help better capture the domain for which the lexicon is being adapted, by enabling the discovery of relevant concepts and semantic relations between terms. Capturing the semantic relationships among terms helps understand the variety of contexts in which terms may be influencing each other.

3 Semantic Approach to Sentiment Lexicon Adaptation

In this section we present our semantic-based approach for sentiment lexicon adaptation. As previously mentioned, the proposed approach extends our previous context-based adaptation model [18] by using conceptual semantics to enrich the context or domain in which the words are used with the aim of enabling a better interpretation of this context. As such, in the following section we briefly describe our general context-based lexicon adaptation model before presenting our new semantic enrichment of this model.

3.1 Context-based Sentiment Lexicon Adaptation

The pipeline of our context-based lexicon adaptation model consists of two main steps, as depicted in Figure 1(a). First, given a tweet collection and a general-purpose sentiment lexicon, our approach detects the context of each word in the tweet collection and uses it to extract the word's contextual sentiment. Secondly, a set of rules are applied to amend the prior sentiment of terms in the lexicon based on their corresponding contextual sentiment. Both steps are further detailed in Sections 3.1.1 and 3.1.2. The semantic enrichment of this pipeline is described in Section 3.2.



(a) Context-based Adaptation

(b) Adaptation with Semantic Relation Adjustment

Fig. 1: Pipelines for (a) Context-based Adaptation Model, and (b) Context-based Adaptation Model with Semantic Relations Adjustment

3.1.1 Word's Contextual Sentiment

The first step in our pipeline is to extract the contextual sentiment of terms (i.e., sentiment extracted based on a word's context) in a given tweet collection. This step consists of: (i) capturing the context in which the word occurs, and (ii) computing the word's contextual sentiment. A common method for capturing the word's context is by looking at its co-occurrence patterns with other terms in the text. The underlying principle behind this method comes from the distributional semantic hypothesis:[3] *words that are used and occur in the same contexts tend to purport similar meanings* [24]. For example, the word "`great`", when occurs in the context "`smile`", denotes a different meaning than when it occurs within the context "`pain`" and "`loss`". Such context variations of the word often affect its sentiment: "`great`" with "`smile`" indicates a positive sentiment, while "`great`" with "`pain`" indicates a negative one.

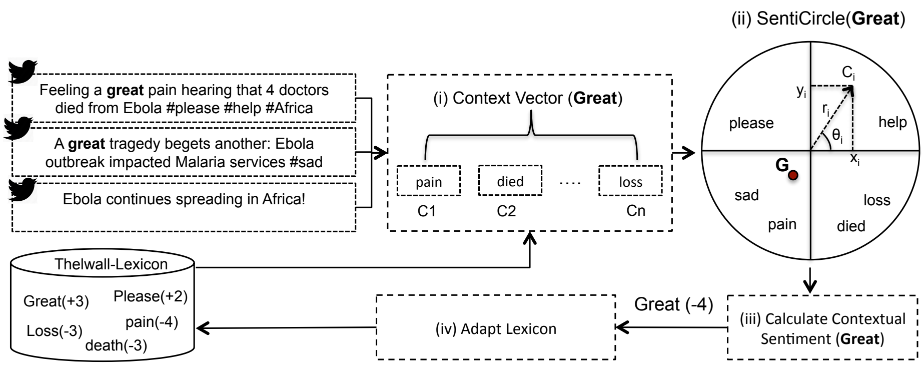


Fig. 2: Illustrative example of extracting the contextual sentiment of the word `Great` from a tweet collection and adapting Thelwall-Lexicon with the new extracted sentiment respectively. Red dot in the SentiCircle represents the geometric median G of the context points in the circle.

Several approaches have been built and used for extracting the words' contextual sentiment following the above principle [23, 10]. In this paper, we use the SentiCircle approach [17], which similarly to other frequency-based approaches, it detects the context of a term from its co-occurrence patterns with other terms in tweets. Figure 2 depicts the representation and extraction of the contextual sentiment of the term "`great`" by the SentiCircle approach. First, given a tweet collection $\mathcal{T}$, the target term m_{great} is represented as a vector $\boldsymbol{c}_{great} = (c_1, c_2, ..., c_n)$ of terms co-occurring with term m in any tweet in $\mathcal{T}$ (e.g., "`pain`", "`loss`", ..., "`death`"). Secondly, the context vector $\boldsymbol{c}_{great}$ is transformed into a 2d circle representation. The center of the circle represents the target term m_{great} and points within the circle denote the context terms of m_{great}. The position (x_{c_i}, y_{c_i}) of each context term $c_i \in \boldsymbol{c}_{great}$ is defined as:

$$x_{c_i} = r_i \cos\theta_i \qquad\qquad y_{c_i} = r_i \sin\theta_i \tag{1}$$

Where the angle θ_i represents the prior sentiment of the context term c_i multiplied by π, and it is obtained from the lexicon to be adapted. The radius r_i represents co-occurrence frequency between c_i and the target term m_{great} and it is computed based on the *TF-IDF* weighting scheme.

[3] Also known as *Statistical Semantics* [25]

Based on the SentiCircle representation, terms with positive prior sentiment are positioned on the upper half of the circle (e.g., `please`) and terms with negative prior sentiment are positioned in the lower half (e.g., `pain`, `loss`). Term co-concurrence determines the distance (i.e., radius) of these terms with respect to the origin. Thirdly, the geometric median G of the SentiCircle of "`Great`" is computed (-4 in our example), which constitutes the contextual sentiment of the term.[4]

3.1.2 Rules for Lexicon Adaptation

The second step in our pipeline is to adapt the prior sentiment of terms in a given sentiment lexicon based on the terms' contextual sentiment information extracted in the previous step. To this end, we propose a general rule-based method to decide on the new sentiment of terms in the lexicon. In the following we give a formal definition of the general purpose sentiment lexicon and its properties, and explain how our proposed method functions on it accordingly.

General-purpose sentiment lexicon: is a set of terms $\mathcal{L} = \{t_1, t_2, ..., t_n\}$ of fixed size n. Each term $t \in \mathcal{L}$ is coupled with a prior sentiment score that is often a numerical value $prior_m \in [-\lambda, -\delta, \delta, \lambda]$, denoting the sentiment orientation and strength of t. In particular, t is *positive* if $prior_t \in (\delta, \lambda]$, *negative* if $prior_t \in [-\lambda, -\delta)$, and *neutral* if $prior_t \in [-\delta, \delta]$. $|\lambda|$ is the maximum sentiment strength that a term can have. The closer the $prior_t$ is to λ the higher the sentiment strength is. $|\delta|$ defines the boundaries of the neutral sentiment range. The values of both, λ and δ depend on the specifications of the studied sentiment lexicon and are defined at the design/construction phase of the lexicon (see section 4).

Lexicon adaptation rules: our proposed method uses a set of 4 antecedent-consequent rules (Table 1) to decide how to update the prior sentiment of a term ($prior_t$) in a given lexicon with respect to its contextual sentiment ($contextual_t$). As noted in Table 1, these rules are divided into (i) **Updating Rules** for updating only the existing terms in the lexicon, and (ii) **Expanding Rules** for expanding the lexicon with new opinionated terms.

Id	Antecedents	Consequent
Updating Rules (Same Sentiment Orientations)		
1	$(\lvert contextual_t\rvert > \lvert prior_t\rvert) \wedge (\lvert contextual_t\rvert > \theta)$	$prior_t = \begin{bmatrix} prior_t + \alpha; prior_t > 0 \\ prior_t - \alpha; prior_t < 0 \end{bmatrix}$
Updating Rules (Different Sentiment Orientations)		
2	$(\lvert contextual_t\rvert > \theta) \wedge (\lvert prior_t\rvert \leqslant \theta)$	$prior_t = \begin{bmatrix} \alpha; prior_t < 0 \\ -\alpha; prior_t > 0 \end{bmatrix}$
3	$(\lvert contextual_t\rvert > \theta) \wedge (\lvert prior_t\rvert > \theta)$	$prior_t = \begin{bmatrix} prior_t - \alpha; prior_t > 0 \\ prior_t + \alpha; prior_t < 0 \end{bmatrix}$
Expanding Rule		
4	$term(t) \notin lexicon(\mathcal{L})$	$(prior_t = contextual_t) \wedge AddTerm(t, \mathcal{L})$

Table 1: Adaptation rules for sentiment lexicons, where $AddTerm(t, mathcalL)$: add term t to lexicon $\mathcal{L}$.

The notion behind the proposed rules is rather simple: *For a given term $t \in \mathcal{L}$, check how strong/weak the contextual sentiment ($contextual_t$) is and how strong/weak the prior sentiment ($prior_t$) is $\rightarrow$ update $prior_t$ in the lexicon accordingly*. As mentioned earlier, $contextual_t$ is obtained as described in Section 3.1.1 and its value range $[-\lambda, \lambda]$.

[4] We refer the reader to the body of [17] for more details about the SentiCircle approach.

The threshold θ is computed as $\theta = |\lambda|/2$ and it is used to determine how *strong/weak* the sentiment of the term is. If the term does not exist in the lexicon, we add it to the lexicon with its corresponding contextual sentiment.

In Thelwall-Lexicon [20], as will be explained in Section 4, $|\lambda| = 5$ and $|\delta| = 1$, i.e., the prior sentiment for the terms in this lexicon is between $[-5, +5]$, and the neutral sentiment range is between $[-1, 1]$. The value of θ is set up to 3.[5]

3.2 Semantic Enrichment for Context-based Lexicon Adaptation

In Section 3.1 we showed our proposed method to adapt sentiment lexicons based on the contextual sentiment of terms in a given collection of tweets. However, relying on the context only for detecting terms sentiment might be insufficient. This often happens either due to the lack of context in tweets, or because the sentiment of a term may be conveyed via its conceptual semantics rather than by its context [4].

In this section we propose enriching our original context-based adaptation model, described in the previous section, with the conceptual semantics of words in tweets. To this end, we propose adjusting the contextual correlation between two co-occurring named-entities in tweets based on the semantic relations between them. We refer to this model as the *Semantically-adjusted Relations Model*.

3.2.1 Semantically-adjusted Relations Model

Using the distributional semantic hypothesis, our context-based approach assigns a stronger relation to words that tend to co-occur more frequently in same context. However the document collection may represent only a partial view of the contexts in which two words my co-occur together. For example, in the GASP Twitter dataset around the dialogue for earth gas prices[1], the entities *Barack Obama* and *Texas* tend to appear together and therefore have a strong contextual relation. However, these two entities are related within a high number of different contexts. Figure 3 shows a small sample of the different semantic contexts that link the two previous entities. These contexts include Barack Obama's birth place, his candidatures and his duties as president.

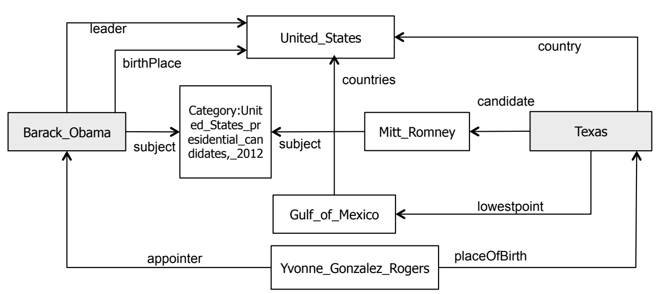


Fig. 3: Example for sentiment relations between the entities *Barack Obama* and *Texas* with a path length of ≤ 3

To capture the variety of contexts in which two terms can potentially appear together we compute the number of relations between these two terms in DBpedia by using the

[5] Since Thelwall-Lexicon uses discrete and not continuous values for priors, θ is rounded up to the nearest integer value to match the annotation format of Thelwall-Lexicon

approach proposed by Pirro [14]. Our assumption is that the strength of the contextual relation between two terms, captured by their co-occurrence within the document collection, should be modified according to the number of contexts in which these terms can potentially appear together. The smaller the number of contexts, the stronger the contextual relation should be.

Based on the above assumption we propose adjusting the strength of the contextual relations between terms, captured by the context-based model, by using the semantic relations between them. To this end, we add two additional steps to the original pipeline (see Figure 1:c): *semantic relation extraction* and *semantic relation adjustment*. These two steps are further described below.

1. Semantic relation extraction: This step extracts the sets of semantic relations for every pair of named entities co-occurring together in the tweets. For the purpose of our study we extract semantic relations using the approach proposed by Pirro [14] over DBpedia, since DBpedia is a large generic knowledge graph which captures a high variety of relations between terms. To extract the set of relations between two name entities this approach takes as input the identifiers (i.e., URIs) of the source entity e_s, the target entity e_t and an integer value K that determines the maximum path length of the relations between the two named entities. The output is a set of SPARQL queries that enable the retrieval of paths of length at most K connecting e_s and e_t. Note that in order to extract all the paths, all the combinations of ingoing/outgoing edges must be considered. Following our previous example, if we were interested in finding paths of length $K <= 2$ connecting $e_s = Obama$ and $e_t = Texas$ our approach will consider the following set of SPARQL queries:

```
SELECT * WHERE {:Obama ?p1 :Texas}
SELECT * WHERE {:Texas ?p1 :Obama}
SELECT * WHERE {:Obama ?p1 ?n1. ?n1 ?p2 :Texas}
SELECT * WHERE {:Obama ?p1 ?n1. :Texas ?p2 ?n1}
SELECT * WHERE {?n1 ?p1 :Obama. :Texas ?p2 ?n1}
SELECT * WHERE {?n1 ?p1 :Obama. ?n1 ?p2 :Texas}
```

As it can be observed, the first two queries consider paths of length one. Since a path may exist in two directions, two queries are required. The retrieval of paths of length 2 requires 4 queries. In general, given a value K, to retrieve paths of length K, 2^k queries are required.

2. Semantic relation adjustment: Now we have for every pair of named entities (e_s, e_t) co-occurring together in the tweets, a set $\mathcal{R}_{(e_s,e_t)} = \{p1, p2, ..., p_N\}$ of paths of size N, representing the semantic relations between e_s and e_t.

As mentioned earlier, our goal behind enriching the context-based model with semantic relations is to adjust the strength of the contextual relation between e_s and e_t based the number of semantic relations (paths) between them. To this end, we construct the SentiCircle S_{e_s} of the source entity e_s, as depicted in Figure 4. Since both entities co-occur together in tweets, the target entity e_t is positioned in the SentiCircle S_{e_s} with a radius r_t representing the strength of the contextual relation between e_s and e_t, as described in Section 3.1.1. Therefore, the task of adjusting the contextual relations between e_s and e_t breaks down into altering the value of r_t as follows:

$$r'_t = r_t + \left[\frac{N}{M}(1 - r_t)\right] \tag{2}$$

Where N is the number of the semantic paths between e_s and e_t extracted in the previous step, M is the maximum number of paths extracted for a pair of entities in the Twitter dataset, and r'_t is the new radius of entity e_t after adjustment.

As can be noted, the above equation modifies the value of r_t based on the number of paths between e_s and e_t. The smaller the number of paths is, the stronger the contextual relation should be, and thereby the higher the value of r'_t is.

Note that the enrichment by semantic relations in this model is done in two iterations of the adaptation process. Specifically, in the first iteration the sentiment lexicon is adapted using the original context-based model (Figure 1:a). In the second iteration the semantically-adjusted relation model is applied on the adapted lexicon, where the semantic-relation adjustment takes place. Adaptation in the first iteration allows us to capture the contextual relations of entities within tweets and assign them a sentiment value. Note that sentiment lexicons are generic and most of the tweet entities (e.g., `Obama`, `Texas`) will not appear in these lexicons. By relying on one iteration of adaptation only, an entity will have little impact on the contextual sentiment of other entities since entities don't generally have any initial sentiment score within the lexicon to be adapted. Hence, a second iteration of adaptation is required in order to detect the sentiment of entities that do not occur in the lexicon, and maximise the impact of the semantic relation adjustment in our models.

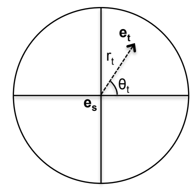


Fig. 4: SentiCircle of entity e_s showing the strength of the contextual relation between e_s and e_t, represented by the radius r_t

4 Experimental Setup

In this section, we report the results obtained from using both, the context-based model (Section 3.1) and the semantically-adjusted relations model (Section 3.2.1) for adapting sentiment lexicons. Evaluation of both models requires the selection of (i) the sentiment lexicon to be adapted, (ii) the context (Twitter datasets) for which the lexicon will be adapted, (iii) the different configurations for adapting the lexicon and, (iv) the semantic information used for the semantic adaptation models:

(i) Sentiment Lexicon For the evaluation we choose to adapt the state-of-the-art sentiment lexicon for social media; Thelwall-Lexicon [21, 20]. Thelwall-Lexicon is a general purpose sentiment lexicon specifically designed to function on social media data. It consists of 2546 terms coupled with values between -5 (very negative) and +5 (very positive), defining their sentiment orientation and strength. Terms in the lexicon are grouped into three subsets of 1919 negative terms ($prior_t \in$[-2,-5]), 398 positive terms ($prior_t \in$[2,5]) and 229 neutral terms ($prior_t \in${-1,1}). Based on the aforementioned specifications, the parameters in our proposed adaptation method (Section 3.1.2) are set as: $|\lambda| = 5$, $|\delta| = 1$, $|\theta| = 3$ and $|\alpha| = 1$.

(ii) Evaluation Datasets To assess the performance of our lexicon adaptation method we require the use of datasets annotated with sentiment labels. for this work we selected three evaluation datasets often used in the literature of sentiment analysis (SemEval, WAB and GASP) [16]. These datasets differ in their sizes and topical focus. Numbers of positive and negative tweets within these datasets are summarised in Table 2.

Dataset	**Tweets**	**#Negative**	**#Positive**	**#Unigrams**
Semeval Dataset (SemEval) [12]	7520	2178	5342	22340
The Dialogue Earth Weather Dataset (WAB) [1]	5482	2577	2905	10917
The Dialogue Earth Gas Prices Dataset (GASP) [1]	6032	4999	1033	12868

Table 2: Twitter datasets used for evaluation. Details on how these datasets were constructed and annotated are provided in [16].

(iii) Configurations of the Lexicon Adaptation Models We test both, the context-based and semantic-adjusted relations adaptations models under three different configurations: (1) *Lexicon Update (LU)*: The lexicon is adapted only by updating the prior sentiment of existing terms, (2) *Lexicon Expand (LE)*: The lexicon is adapted only by adding new opinionated terms, and (3) *Lexicon Update and Expand (LUE)*: The lexicon is adapted by adding new opinionated terms and by updating the prior sentiment of existing terms.

Dataset	No. of Entities	numRelations	minPath	maxPath	AveragePath
SemEval	2,824	1,011,422	1	3	2.82
GASP	685	811,741	1	3	2.95
WAB	750	796,021	1	3	2.92

Table 3: Amount of relations and path lengths extracted for each dataset

(iii) Extracted Semantics As mentioned in Section 3.2.1 semantic enrichment of our adaptation model relies on the strength and diversity of the semantic relations between entities in tweets. Table 3 shows the number of named entities extracted from each Twitter dataset along with the semantic relations between them. This table also includes the minimum, maximum and average path length among all the extracted relations. A maximum path length of 3 was consider for our experiments.

5 Evaluation Results

In this section, we report the results obtained from using the different adaptations of Thelwall-Lexicon to compute tweet-level sentiment detection. To compute sentiment, we use the approach proposed by Thelwall [20], where a tweet is considered positive if its aggregated positive sentiment strength (i.e., the sentiment strength obtained by considering the sentiment weights of all words in the tweet) is 1.5 times higher than the aggregated negative one, and vice versa. Our baselines for comparison are the original version of Thelwall-Lexicon.

Results in all experiments are computed using 10-fold cross validation over 30 runs of different random splits of the data to test their significance. The null hypothesis to be tested is that for a given dataset, the baseline lexicons and the lexicons adapted by our models will have the same performance. We test this hypothesis using the *Paired T-Test* since it determines the mean of the changes in performance, and reports whether this mean of the differences is statistically significant. Note that all the results in F1-measure reported in this section are statistically significant with $\rho < 0.001$.

5.1 Results of Context-based Lexicon Adaptation

The first task in our evaluation is to assess the effectiveness of our context-based adaptation model. Table 4 shows the results of binary sentiment detection of tweets performed on the three evaluation datasets using (i) the original Thelwall-Lexicon (*Original*), (iii) Thelwall-Lexicon adapted under the update setting (LU), (iv) Thelwall-Lexicon adapted under the expand setting (LE), and (v) Thelwall-Lexicon adapted under the update and expand setting (LUE). The table reports accuracy and three sets of precision (P), recall

(R), and F1-measure (F1), one for positive sentiment identification, one for negative sentiment identification, and the third showing the average of the two.

Dataset	Lexicon	Accuracy	Negative Sentiment			Positive Sentiment			Average		
			P	R	F1	P	R	F1	P	R	F1
Average	Original	73.49	71.90	79.09	73.88	66.30	64.23	63.77	69.10	71.66	68.83
	LU	**76.34**	75.29	76.76	75.29	66.06	68.80	66.70	**70.68**	**72.78**	**71.00**
	LE	73.50	72.03	78.42	73.74	65.99	64.67	63.90	69.01	71.55	68.82
	LUE	76.22	75.13	76.54	75.12	65.94	68.68	66.59	70.53	72.61	70.85

Table 4: Results obtained from adapting Thelwall-Lexicon on three datasets using the context-based adaptation model. Bold=highest performance. LU=Lexicon Update, LE=Lexicon Expand, and LUE=Lexicon Update and Expand.

Overall, the average performance across the three datasets shows that the improvement of the adapted LU and LUE lexicons over the original lexicon 3.9% in accuracy, and 3.2% in F1. On the other hand, the LE lexicon gives negligible performance improvements over the original lexicon.

5.2 Results of the Semantically-adjusted Relations Model

The second step in our evaluation is to test the performance of the adapted lexicons by the semantically-adjusted relations model (Section 3.2.1). The lower part of Table 5 lists the average results across the three datasets for the adapted lexicon under the update setting (SRU), the expand setting (SRE), and the update and expand setting (SRUE).

Model	Lexicon	Accuracy	Negative Sentiment			Positive Sentiment			Average		
			P	R	F1	P	R	F1	P	R	F1
Baselines	Original	73.49	71.90	79.09	73.88	66.30	64.23	63.77	69.10	71.66	68.83
Semantically-adjusted Relations Model	SRU	76.14	75.99	74.74	74.50	64.78	68.95	66.14	70.39	71.84	70.31
	SRE	**76.66**	77.38	73.62	74.76	64.84	71.42	67.31	**71.11**	**72.52**	**71.03**
	SRUE	76.13	76.03	74.75	74.52	64.78	69.01	66.16	70.41	71.88	70.34

Table 5: Average results across the three datasets of Thelwall-Lexicon adapted by the semantic model. Bold=highest performance.

According to these results in Table 5, we notice that the three semantically adapted lexicons SRU, SRE and SRUE outperform the original lexicon by a large margin. In particular, the lexicon adapted under the expand setting, SRE outperforms both baseline lexicons by 4.14% in accuracy and 3.1% in average F1. The SRU and the SRUE lexicons come next by a performance that is 3.5% and 2.1% higher in accuracy and F1 than the baseline.

Figure 5 shows the the win/loss in accuracy, P, R and average F1 when using semantically-adjusted relations model for lexicon adaptation compared to the context-based model across the three datasets. Here we notice that the *expand setting* the semantically-adjusted relation model boosts the performance substantially, with 4.12% and 3.12% gain in accuracy and F1 respectively. On the other hand, a different performance trend can be noted for both, the *lexicon update setting* and the *lexicon update & expand setting*, where the semantically-adjusted relations model always gives, under these settings, a lower performance on all measures compared to the context-based model.

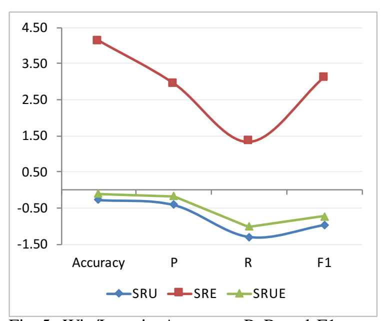


Fig. 5: Win/Loss in Accuracy, P, R and F1 measures of adapting sentiment lexicons by the semantic relations model in comparison with the context-based model.

6 Conclusions and Future Work

Although much research has been done on creating domain-specific sentiment lexicons, very little attention has been giving to the problem of lexicon adaptation in social media, and to the use of semantic information as a resource to perform such adaptations.

This paper proposed a general method to adapt sentiment lexicons based on contextual information, where the domain or context of adaptation is defined by a collection of posts. A semantic enrichment of this method is also proposed where semantic relations between named entities in the text are used to better capture the context for which the lexicon is being adapted.

An evaluation of our proposed method was performed by adapting the state-of-the-art sentiment lexicon for the social web [21] to three different contexts (Twitter datasets) using various configurations of our proposed approach. Results showed that the adapted sentiment lexicons outperformed the baseline method in average by up to 3.9% in accuracy and 3.2% in F1 measure, when used to compute tweet-level polarity detection with context-based adaptation. While enriching the adaptation process with the semantic relations between entities in tweets, yields in 4.12% and 3.1% gain in accuracy and F1 measure in comparison with context-based adaptation.

While this initial results are positive, our research is still far from exploiting the full potential of *conceptual semantics*. For the moment, only the number of semantic relations has been taken into consideration (number of contexts), but not the concrete relations that emerge between two concepts (i.e., the particular contexts in which two words may appear together). As future work we plan to refine the SentiCircle model by adding and removing terms based on the particular contexts in which two words may appear together.

In addition, while DBpedia is a core element of the Linked Open Data (LOD) graph our approach only exploits a very small subsection of the information available in LOD. As future work we plan to extend our relation extraction process so that multiple datasets can be considered and more fine-grained relations, expanding multiple datasets, can be taken into consideration when adapting the lexicons.

Acknowledgment

This work was supported by the EU-FP7 project SENSE4US (grant no. 611242).

References

1. Asiaee T, A., Tepper, M., Banerjee, A., Sapiro, G.: If you are happy and you know it... tweet. In: Proc. 21st ACM conference on Information and knowledge management (2012)
2. Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In: Seventh conference on International Language Resources and Evaluation. Valletta, Malta (2010)
3. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. Journal of Computational Science 2(1), 1–8 (2011)
4. Cambria, E.: An introduction to concept-level sentiment analysis. In: Advances in Soft Computing and Its Applications, pp. 478–483. Springer (2013)
5. Cambria, E., Havasi, C., Hussain, A.: Senticnet 2: A semantic and affective resource for opinion mining and sentiment analysis. In: FLAIRS Conference. pp. 202–207 (2012)

6. Choi, Y., Cardie, C.: Adapting a polarity lexicon using integer linear programming for domain-specific sentiment classification. In: Proc. Conf. Empirical Methods in Natural Language Processing (2009)
7. Hu, X., Tang, J., Gao, H., Liu, H.: Unsupervised sentiment analysis with emotional signals. In: Proceedings of the 22nd World Wide Web conf (2013)
8. Kaewpitakkun, Y., Shirai, K., Mohd, M.: Sentiment lexicon interpolation and polarity estimation of objective and out-of-vocabulary words to improve sentiment classification on microblogging. In: Proc. 28th Pacific Asia Conf. Language, Information and Computing. Department of Linguistics, Faculty of Arts, Chulalongkorn University, Thailand (2014)
9. Li, F., Pan, S.J., Jin, O., Yang, Q., Zhu, X.: Cross-domain co-extraction of sentiment and topic lexicons. In: Proc. 50th Annual Meeting of Association for Computational Linguistics (2012)
10. Lin, C., He, Y., Everson, R., Ruger, S.: Weakly supervised joint sentiment-topic detection from text. Knowledge and Data Engineering, IEEE Transactions on 24(6), 1134–1145 (2012)
11. Lu, Y., Castellanos, M., Dayal, U., Zhai, C.: Automatic construction of a context-aware sentiment lexicon: an optimization approach. In: Proc. Int. Conf. World Wide Web (2011)
12. Nakov, P., Rosenthal, S., Kozareva, Z., Stoyanov, V., Ritter, A., Wilson, T.: Semeval-2013 task 2: Sentiment analysis in twitter. In: Proc. 7th ACL Workshop on Semantic Evaluation. (2013)
13. Nielsen, F.A.: A new anew: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903 (2011)
14. Pirrò, G.: Explaining and suggesting relatedness in knowledge graphs. In: The Semantic Web-ISWC 2015, pp. 622–639. Springer (2015)
15. Qiu, G., Liu, B., Bu, J., Chen, C.: Expanding domain sentiment lexicon through double propagation. In: IJCAI. vol. 9, pp. 1199–1204 (2009)
16. Saif, H., Fernandez, M., He, Y., Alani, H.: Evaluation datasets for twitter sentiment analysis. In: Proceedings, 1st ESSEM Workshop. Turin, Italy (2013)
17. Saif, H., Fernandez, M., He, Y., Alani, H.: SentiCircles for contextual and conceptual semantic sentiment analysis of twitter. In: Proc. 11th Extended Semantic Web Conf. (ESWC). Crete, Greece (2014)
18. Saif, H., He, Y., Fernandez, M., Alani, H.: Adapting sentiment lexicons using contextual semantics for sentiment analysis of twitter. In: ESWC 2014 Satellite Events (2014)
19. Scheible, C., Schütze, H.: Bootstrapping sentiment labels for unannotated documents with polarity pagerank. In: LREC. pp. 1230–1234 (2012)
20. Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment strength detection for the social web. J. American Society for Information Science and Technology 63(1), 163–173 (2012)
21. Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., Kappas, A.: Sentiment strength detection in short informal text. J. American Society for Info. Science and Technology 61(12) (2010)
22. Turney, P.: Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL'02). Philadelphia, Pennsylvania (2002)
23. Turney, P., Littman, M.: Measuring praise and criticism: Inference of semantic orientation from association. ACM Transactions on Information Systems 21, 315–346 (2003)
24. Turney, P.D., Pantel, P., et al.: From frequency to meaning: Vector space models of semantics. Journal of artificial intelligence research 37(1), 141–188 (2010)
25. Weaver, W.: Translation. Machine translation of languages 14, 15–23 (1955)
26. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing. Vancouver, British Columbia, Canada (2005)