Multilingual Sentiment Analysis: From Formal to Informal and Scarce Resource Languages

Siaw Ling Lo[1], Erik Cambria[2]*, Raymond Chiong[1], David Cornforth[1]

[1]School of Design, Communication and Information Technology, The University of Newcastle, Callaghan, NSW 2308, Australia

[2]School of Computer Engineering, Nanyang Technological University, 639798 Singapore

*Corresponding author
E-mail: cambria@ntu.edu.sg
Phone: (+65) 6790 4328

Abstract

The ability to analyse online user-generated content related to sentiments (e.g., thoughts and opinions) on products or policies has become a de-facto skillset for many companies and organisations. Besides the challenge of understanding formal textual content, it is also necessary to take into consideration the informal and mixed linguistic nature of online social media languages, which are often coupled with localised slang as a way to express 'true' feelings. Due to the multilingual nature of social media data, analysis based on a single official language may carry the risk of not capturing the overall sentiment of online content. While efforts have been made to understand multilingual sentiment analysis based on a range of informal languages, no significant electronic resource has been built for these localised languages. This paper reviews the various current approaches and tools used for multilingual sentiment analysis, identifies challenges along this line of research, and provides several recommendations including a framework that is particularly applicable to dealing with scarce resource languages.

Keywords: multilingual analysis, sentiment analysis, scarce resource languages, social media

1. Introduction

Sentiment analysis has been a popular research area over the past few years. It is gaining even more attention with the prevalence of social media usage, where netizens freely and openly express their views and opinions about anything; be it a product, a policy or even a picture. Although these opinions are valuable for understanding the concerns and issues on the ground, it remains a challenge to fully decipher the message and context of online user-generated content. This is mainly due to a few key issues, such as sentence parsing, named entity recognition, anaphora resolution and concept disambiguation. It is essential to comprehend the subject and topic of any content before discerning the sentiment expressed (e.g., positive or negative). To make the matter more complicated, online sharing or social media content is known to be noisy and often mixed with linguistic variations. It is thus not surprising that sentiment analysis continues to be one of the main analytics research domains given its many challenges but also promises.

Sentiment analysis for a language is usually dependent on manually or semi-automatically constructed lexicons [1], [2], found in dictionaries or corpora [3]. The availability of these resources enables the creation of rule-based sentiment analysis or the construction of training data for classification tasks. Despite the fact that English remains the main language used in various research studies in this area (e.g., see [4], [5]), there are also efforts in creating subjectivity resources for other formal languages such as Japanese [6], Chinese [7] and German [2]. However, since creating lexical or corpus resources for a new language can be very time-consuming and resource intensive, most of the multilingual sentiment analyses on other languages [3], [8] have been relying on some available English knowledge base, such as SentiWordNet [9].

While increasing effort has been made in creating resources for other formal languages, there are not many resources available when it comes to languages that are not commonly used in official communication or formal news reporting due to their informal and evolving nature. These languages often evolve from a main national language, such as English, and are broadly used by a local community in daily conversation both in the physical and online world. With the popularity of social media and the freedom of expression it affords, languages with localised expressions or variants of formal languages are becoming widespread in the online environment. In addition, it is not uncommon to see a few languages being mixed to form a unique language in a multicultural society. One such example is Singlish, the colloquial Singaporean English that has incorporated elements of some Chinese dialects and the Malay language [10]. To fully understand the sentiments in this sort of languages, it is essential to analyse them alongside other formal languages. The aim of this paper is to review sentiment analysis research in a multilingual setting, by considering not just formal but also informal and scarce resource languages used on social media, especially variants of the English language. It is of interest to examine current approaches and tools used in multilingual sentiment analysis, so that challenges can be identified and recommendations can be provided.

By scarce resource languages, we refer to those with just a basic dictionary available and/or lacking of developed text processing resources (such as a translation engine). The various English variants widely used on social media belong to this category. In this paper, we first assess a range of current multilingual sentiment analysis studies based on the resources used, in terms of whether a lexicon, a corpus, a translation machine or a translator is applied, before sentiment analysis research carried out on social media is reviewed. It is important to examine current approaches used in analysing social media data, given that the world is at present dominated by this kind of data. Most of the social media messages would be written in an informal manner, with linguistic variations that require different considerations compared to analysing formal reviews or news corpora that would typically consist of a single official language. Social media data analysis can be treated as understanding another 'new' language with limited resources. Here, we handle sentiment analysis of social media data separately with respect to other scarce resource languages, as the majority of research studies on scarce resource languages have been focused on a single language.

In the next section, we describe current approaches used in multilingual sentiment analysis studies. We then cover other types of studies on multilingual sentiment analysis, and list resources available for different languages. This is followed by reviewing sentiment analysis research carried out on social media, before touching on research done on other scarce resource languages. After that, we put forth the challenges identified and recommendations to overcome these challenges. The recommendations include some proposed solutions and a hybrid framework for dealing with scarce resource languages. Finally, we conclude the paper.

2. Current approaches used for multilingual sentiment analysis

There are mainly two approaches in sentiment analysis – subjectivity and polarity detection. Subjectivity detection is about understanding if the content contains personal views and opinions as opposed to factual information. Often, these subjective expressions are due to culture or experience of a person or community and hence, can be very 'localised' and specific to a society. As a result, subjectivity is usually studied before detailed sentiment analysis is done, since it is essential to filter out factual content to have a better understanding of issues that are shared among netizens. Polarity detection, on the other hand, is about studying subjectivity with different polarities, intensities or rankings. Some polarity analysis studies regarded an opinion as either highly positive, positive, negative or highly negative [4], while others [11] worked on human emotion such as joy or anger.

Most subjectivity and polarity analysis studies have limited themselves to English, but with the increasing popularity of online social media worldwide, it is no longer sufficient to deal with only English language content. In fact, only 28.6% of the Internet users speak English[1]. It is thus essential to explore or build resources and tools in languages other than English. Moreover, Asia now has the most Internet users (48.2%); followed by Europe (18%)[2]. As a result, there is a growing need to work on languages such as Chinese and Japanese. Multilingual subjectivity and polarity analysis research has become more widespread, and languages that have been studied include Chinese [7], [12], [13], Japanese [14], German, Spanish, French, Italian [15], Swedish [16], Arabic [17] and Romanian [3].

This review paper will look at the various multilingual approaches taken in the areas of subjectivity and polarity analysis, and assess how these approaches can be applied to a scarce resource language. The general approaches for both subjectivity and polarity analyses on multilingual studies are lexicon, corpus or translator-based, although there are also approaches that move towards research based on concepts and sentics. Sentic computing [18] incorporates common-sense reasoning to specify affective information associated with real-world objects, actions, events and people. The various multilingual sentiment analysis approaches used can be found in Table 1, and their corresponding lexicon, corpus or dataset is listed in Table 2.

2.1 Subjectivity analysis

Mihalcea et al. [3] investigated both lexicon and corpus-based approaches for multilingual subjectivity analysis (subjectivity vs. objectivity). Their lexicon-based approach uses a lemmatised form of English terms from OpinionFinder [5], an English subjectivity analysis system, and translates them into Romanian terms using two bilingual dictionaries. They then built a rule-based subjectivity classifier using the lexicon. The subjectivity precision of the classifier was shown to be good, although its recall was low. Within the same study, corpus-based sentence level subjectivity analysis was conducted based on a parallel corpus consisting of 107 documents from the SemCor corpus [19]. A Naïve Bayes (NB) [20] classifier was used on the Romanian training dataset, where the

[1] http://www.internetworldstats.com/stats7.htm
[2] http://www.internetworldstats.com/stats.htm

annotations were projected from two OpinionFinder [5] classifiers. While the highest precision for subjective classification was obtained with the rule-based classifier using the generated lexicon, the overall best F-measure result of 67.85 was produced by the NB-based statistical machine learning approach.

Ahmad et al. [21] used a local grammar approach to extract sentiment-bearing phrases within a multilingual framework (English, Arabic, and Chinese). As their focus was on sentiment analysis of financial news streams, domain-specific keywords were selected by comparing the distribution of words in domain-specific documents to the distribution of words in a general language corpus. Words from domain-specific documents found to be asymmetric with the general corpus were assigned as keywords. These keywords, together with their local grammar patterns, were used to extract sentiment-bearing phrases. Their experimental results showed that the local grammar patterns in all three languages considered, i.e., English, Arabic and Chinese, can be used to extract sentiment-bearing phrases. This observation is important, as it demonstrates that domain-specific keywords can transcend different language typologies (Indo-European -> Sino-Asiatic -> Semitic). Their manual evaluation found that the accuracy of their approach is within the 60-75% range.

It is worth noting that the approaches listed above do not determine the polarity of content but focus on construction and detection of words or phrases containing subjectivity notions. Although subjectivity analysis does not apply directly to sentiment analysis or opinion mining, it is often the first step towards improving sentiment classification results [4]. It has been shown that distinguishing subjective versus objective instances is often more challenging than the subsequent polarity classification [22], [23].

2.2 Polarity analysis

There are different granularities of polarity analysis. Basic analysis involves classifying the expressed opinion of given text (e.g., at the aspect, sentence or document level) as being positive, negative or neutral. More advanced analysis deals with classification at the emotion or affective level, where different emotion states such as “joy”, “angry” and so on are recognised. This review paper concentrates on the methods/approaches taken with regards to whether a lexicon, corpus or translation engine is used, and hence both the analyses of positive/negative expressions and various emotion states are considered.

In contrast to subjectivity analysis, polarity analysis is not limited to lexicon- or corpus-based approaches. While lexical resources are still used to detect the polarity in text, machine-learning approaches are more common in this type of analysis. In addition, machine translation engines or translators are often used in conjunction with various English knowledge bases. Concept-based resources such as SenticNet [11] are also used for multilingual sentiment analysis.

2.2.1 Lexicon and machine learning-based polarity analysis

One of the first studies on multilingual polarity analysis can be found in the work of Yao et al. [24], in which they proposed a method to determine sentiment orientation of Chinese words by using a bilingual lexicon. Their method uses the occurrence of English sentiment words from an interpreted Chinese word to predict the sentiment orientation of the Chinese word. This is achieved through the calculation of the sentiment vector from the English word sequence followed by classification based on the Support Vector Machine (SVM) [25] and C4.5 [26]. The best accuracy obtained in predicting the sentiment orientation of a Chinese word is above 90%, when support vectors that do not contain any polarity words are eliminated from the classification.

Kim and Hovy [2] utilised a lexical database, i.e., WordNet [27], and three sets of manually annotated positive, negative and neutral words to build a word sentiment classifier for detecting opinions in emails. Since their opinion-bearing words are in English and the target system is in German, a statistical word alignment technique, GIZA++ [28], is used on a parallel European Parliament corpus to acquire word pairs in German-English and English-German. These word pairs are then used to build a German opinion analysis system using the English opinion-bearing words without a translation system. The precision obtained is 72% for positive emails and the recall is 80% for negative emails, but the recall and precision values for positive and negative emails, respectively, are low.

In a different study, Rosell and Kann [16] constructed a Swedish general purpose polarity lexicon with a graph-based random walk approach. Using the People's Dictionary of Synonyms [29], they extracted a large amount of polarity terms from a small set of seed words through mapping from a bilingual dictionary of English and Swedish languages. Their random walk approach takes into consideration the synonymity and path length in calculating the mean polarity value of words. Some examples of words with their polarity values have been presented.

Another lexical resource for sentiment analysis in English is SentiWordNet [30], used by Denecke [9] to detect the polarity of a document within a multilingual framework. The classification here is based on three classifiers: LingPipe Classifier[3], SentiWordNet Classifier with classification rules, and SentiWordNet Classifier with machine learning. These classifiers were trained using the annotated movie reviews dataset from LingPipe but evaluated on two different testing datasets. The first dataset was generated from the multi-perspective question answering (MPQA) [31] corpus, with 250 positive and 250 negative sentences selected at random. The second dataset was based on German movie reviews selected from Amazon.de, with 100 positive and 100 negative reviews translated to English. Results from the study show that the machine-learning based SentiWordNet Classifier has achieved the best accuracy of 66% for German movie reviews, while the other two classifiers have similar accuracies of around 52% for English and 58% for German documents. In addition, the results suggest that the accuracy of the different methods does not depend on the processed language.

Wan [32] used the English sentiment lexicon from OpinionFinder [5] for Chinese sentiment analysis by employing machine translation and ensemble techniques. Experimental results show that using an ensemble of Chinese lexicons with English reviews translated by both Google Translate and Yahoo BabelFish can achieve an accuracy of 0.854. Wan further extended the lexicon-based approach to a corpus-based one via a co-training method using two-way translation [8], so that the English and Chinese features can be considered as two independent views of the classification problem. Labelled English reviews are used to create labelled Chinese reviews through translation. The unlabelled Chinese reviews are paired with the labelled Chinese reviews (translated from English reviews) for the first training dataset. The second training dataset is from the translated unlabelled English reviews (derived from Chinese reviews) paired with initially labelled English reviews. The classifiers from the two training datasets are then combined into a single sentiment classifier through a co-

[3] http://alias-i.com/lingpipe/index.html

training process. The co-training approach achieves the best accuracy of 0.775 and 0.79 for English and Chinese classifiers, respectively. This co-training approach is useful in the absence of a parallel corpus, which is covered in the next section.

2.2.2 Parallel corpus-based polarity analysis

Another type of polarity analysis is to use parallel corpora to learn language characteristics without the need of using a translation machine or translator. Meng et al. [33] built a generative cross-lingual mixture model (CLMM) to leverage unlabelled bilingual parallel data. The CLMM utilises words from a parallel corpus to learn about word polarity. It expands the vocabulary through maximising the likelihood of and estimating word-generation probabilities for words not seen in the labelled data but present in the parallel corpus. It is shown that the accuracy of classification results using only English labelled data is 71% but the accuracy improves to 83% when both English and Chinese labelled data are used. The initial lower accuracy is probably due to the limited vocabulary coverage in machine translated data and hence the usage of the parallel bilingual corpus improves the classification results by learning previously unseen sentiment words from the large unlabelled data.

Lu et al. [34] adopted a maximum entropy-based approach to jointly learn two monolingual sentiment classifiers. Their focus is to simultaneously improve the performance of sentiment classification in a pair of languages – English and Chinese – by relying on sentiment-labelled data in each language as well as unlabelled parallel text for the language pair. It is reported that the proposed approach is able to outperform the monolingual baselines and improve the accuracy for both languages by 3.44%-8.12%, with the best accuracy scored at 83.71% by the English classifier using the NTCIR parallel corpora [35], [36].

2.2.3 Corpus and machine learning-based polarity analysis

In contrast to the parallel corpora approach, Prettenhofer and Stein [37] used English as the source language, and German, French and Japanese as target languages, for cross-language topic and sentiment classification. Structural Correspondence Learning (SCL) [38], proposed for domain adaption, was adopted in their study. Unlabelled documents from both languages, together with pivot words or pairs of words that have predictive value, were used to create a map of cross-lingual feature space. It is shown that their approach can reduce the relative error to 59% in sentiment classification as compared to a machine translation baseline.

Boiy and Moens [39] also did not use language translation in their work. Instead, they used three manually annotated languages – English, Dutch and French – to train various machine learning algorithms for classifying if a statement is positive, negative or neutral with regards to a certain entity. They proposed a cascading framework for the three languages, but different negation rules, discourse processing and parsing tools were used for each of the languages. This is mainly due to the different behaviours of the languages and the fact that different machine learning algorithms also work differently. Their results show that an English corpus using unigram features augmented with linguistic features achieves an accuracy of 83%, while Dutch and French texts have lower accuracies of 70% and 68% because of the larger variety of linguistic expressions in the two languages. The best classification results for English, Dutch and French came from Multinomial Naïve Bayes (MNB), SVM and Maximum Entropy classifiers, respectively.

2.2.4 Corpus-based topic modelling polarity analysis

While most of the corpus-based approaches are coupled with either machine translation or parallel corpora to classify the subjectivity or polarity of given text, Boyd-Graber and Resnik [40] developed a generative topic model known as multilingual supervised Latent Dirichlet Allocation (MS-LDA). Their approach jointly models topics that are consistent across languages, and connects them to predict sentiment ratings. MS-LDA is capable of clustering thematically coherent topics together with their sentiments without requiring parallel corpora and machine translation. It is shown that the model is able to make better prediction when a mix of English and German data is used, compared to when German data alone is used. This is interesting, as the approach shows the potential of leveraging another language to improve sentiment analysis classification results.

2.2.5 Cross-lingual and machine translation polarity analysis

Another polarity analysis approach is to use cross-lingual corpora for multilingual sentiment analysis. Cross-language classification uses a source language (often annotated) as the training dataset and another language or the target language as the testing dataset. It is not uncommon to have documents from the training and testing datasets mapped onto non-overlapping regions of the feature space when the domains of both sources are different. Pan et al. [41] utilised an annotated sentiment corpus in English to predict sentiment polarity in Chinese. The approach uses machine translation so that two datasets in the two languages can be created as two independent views. The two views are combined in a matrix factorisation process so that training can be done simultaneously (instead of conducting training using a series of classifiers from a co-training approach). In addition, lexical knowledge is incorporated into the model to improve its accuracy. Three different datasets (i.e., movie, book and music reviews) were tested in the study and the best accuracy of 84% came from the movie reviews dataset.

Similar to Pan et al. [41], Bautin et al. [42] also used lexicons, translators and two types of corpora (i.e., multilingual news streams and parallel corpora) for sentiment analysis and cross-cultural comparison. Their focus was on comparing the diversity of different languages based on a selected entity, e.g. a politician, over a time period, and they emphasised that it is essential to apply normalisation coefficients to minimise the effect of variance in different languages. The Lydia sentiment system [43] was used and certain entities were selected for cross-language sentiment analysis using 10 days of news streams. Entity sentiment (subjectivity and polarity) was calculated for each day based on co-occurring of the entity with sentiment words. Even though machine translation has been used for the study, it is found that the accuracy is largely translator independent. In addition, the results from a news entity frequency correlation study show that English has a significant correlation with the other eight languages investigated, and hence confirm its pivotal role in the multi-language analysis approach.

2.2.6 Translation-based polarity analysis

One of the reasons for using a parallel corpus is due to the language gap and difference in the underlying distribution between the original language and the translated language [8], [33]. While poor performance of multilingual sentiment analysis may be due to the limitation of a machine translation system, Balahur and Turchi [44] conducted extensive evaluation scenarios to show that machine translation systems are mature enough to obtain multilingual data for supervised sentiment analysis. They quantified the effect of translation quality using three different machine translation systems. Various features, algorithms and meta-classifiers were adopted for polarity detection, and they showed that feature representation using Term Frequency – Inverse Document Frequency of unigram and bigram in an SVM with sequential minimal optimisation produces the best result.

Hiroshi et al. [45] also explored a translation-based approach, which includes parsing and pattern discovery for multilingual sentiment analysis. Specifically, they used transfer-based machine translation technology to develop a high-precision sentiment analysis system for the Japanese language by leveraging English sentiment resources to identify relevant sentiment units. The sentiment unit polarity extraction precision was reported to be as high as 89%.

2.2.7 Concept-based polarity analysis

While lexicon, corpus and translator-based approaches or a combination of these approaches have been used extensively for subjectivity and polarity analysis, concept-based techniques are gaining popularity due to their ability to detect subtly expressed sentiments [46]-[48]. SenticNet [11] is a widely used concept-based resource. Xia et al. [49] created a localisation toolkit for SenticNet by implementing a set of concept disambiguation algorithms to discover context. In this toolkit, Google translate is used to do mapping of the English and Chinese languages. Various Chinese resources are also used to discover language-dependent sentiment concepts through translation. They evaluated the toolkit based on the correctly predicted polarity of the root concept, and an agreement rate of 0.901 was achieved based on annotations from two postgraduate students.

2.2.8 Summary

In short, it is observed that multilingual sentiment analysis using a parallel corpus instead of machine translation can improve classification accuracy [33], [34]. On top of that, Lu et al. [34] showed that a natural parallel corpus produces performance gain compared to using pseudo-parallel data from machine translation engines. Having said that, there are other researchers who firmly believe that machine translation technology has matured [44], and that the techniques used in translation [45] can be applied to multilingual sentiment analysis. Both of these approaches, however, do not work well for scarce resource languages, as parallel corpora and translation machines are literally non-existent for this sort of languages, and manual efforts are needed for creating such resources before the approaches reviewed above can be adopted.

Table 1. Multilingual approaches used in subjectivity and polarity studies

*L, C and T at the table header indicate if the approach uses lexicon, corpus or translator-based resources, respectively. The corresponding resources can be found in Table 2 and Table 3.

Approach	**Challenges**	**Language**	**L***	**C***	**T***	**Reference**
Subjectivity						
Bilingual dictionary translation and rule-based classifier	• Due to inflected English words, lemmatised subjective English terms are used to map entries to the bilingual dictionary but this may lose subjectivity	English, Romanian	√			[3]

Approach	Challenges	Language	L*	C*	T*	Reference
	• Ambiguity of word sense and part of speech due to identical entries • Multi-word expressions that cannot match the dictionary entries so word-by-word matching is adopted					
Parallel annotation projection and statistical classifier	• Interpretation of different languages on subjectivity of a sentence due to different opinions of annotators and loss of information in translation • Difficulty in capturing subtle expressions such as irony	English, Romanian		√		[3]
Local grammar pattern discovery and domain specific keywords	• Word sense ambiguity of sentiment words extracted • Grid-based analysis is proposed to cope with multiple news sources and the huge volume	English, Arabic, Chinese		√		[21]
Polarity						
Lexicon-based to build Support Vector (SV) of sentiment words for the SVM and C4.5	• Manual annotation is used for creating sentiment-tagged Chinese words • It is observed that SV with zero elements (no match in all the positive or negative words) should be eliminated to improve the classifier's result	English, Chinese	√			[24]
Lexicon and parallel corpus with a statistical word alignment approach	• Huge manual effort needed for generating a list of sentiment-bearing words based on WordNet • Results show that the approach recognises negative emails better than positive emails	English, German	√	√		[2]
Lexicon with rule-based classifier and machine learning (Simple Logistic Classifier)	• Translation errors or missing translation • Ambiguities and different meanings of a synset in SentiWordNet are not resolved • Limited ability to recognise negative text; may be due to negated structures not considered in the classifier	English, German	√		√	[9]
Lexicon-based random walk approach on synonyms with seed words and a bilingual dictionary	• Heavily dependent on the dictionary of synonyms and the weight derived from the links between the words	English, Swedish	√			[16]
Lexicon-based	• Cross-lingual lexicon translation	English,	√		√	[32]

Approach	Challenges	Language	L*	C*	T*	Reference
with translation	does not work well for Chinese sentiment analysis	Chinese				
Parallel corpus-based via learning sentiment words from the corpus	• Volume and quality of bilingual parallel data is critical to the performance of the method	English, Chinese		√		[33]
Parallel corpus-based using joint training on two monolingual classifiers on the unlabelled corpus	• It is assumed that the perspectives of parallel sentences in the corpus are the same and should have the same sentiment polarity	English Chinese		√		[34]
Corpus-based with domain adaptation of SCL	• It is essential to have a task or domain specific corpus for the approach • The pragmatic correlation of pivot words or word pairs can only work on a domain specific cross-lingual corpus	English, German, French, Japanese		√		[37]
Corpus-based but with aspect focus	• Manually annotate training data in regard to a certain entity • Major cause of errors is the scarcity of training examples with informal languages used on blogs	English, Dutch, French		√		[39]
Corpus-based with MS-LDA	• Various resources are needed as a bridge to link the different corpora • Quality and the amount of corpora are essential for better performance • Mapping that captures the local syntax and meaningful collocations can improve the model	English, German, Chinese		√		[40]
Corpus-based and 2-way translation with co-training	• Inaccuracy of machine translation service causes the difference in feature distribution • Learning curve of classifiers in the co-training approach	English, Chinese		√	√	[8]
Corpus-based and translation with LingPipe classifier	• Limited ability to recognise negative text; may be due to negated structures not considered in the classifier • Frequency of polarity features and subjectivity detection methods are proposed to improve accuracy	English, German		√	√	[9]
Corpus-based with translation and machine learning	• Translation engines or translators need to be available for the target language • Multiple translated data from	English, Spanish, French, German		√	√	[44]

Approach	Challenges	Language	L*	C*	T*	Reference
	various translators proposed in order to minimise the translation error					
Lexicon and corpus-based with translation to create a bi-view non-negative matrix tri-factorisation model	• Domain specific datasets are used in the study; it is not known how the approach performs on general cross-lingual classification • Parameters used have influence on different languages and they need to be set manually; suggested to estimate the parameters via a validation set	English, Chinese	√	√	√	[41]
Lexicon and corpus-based with translation to English to understand the diversity of the different languages	• The availability of translators for the target language • Due to the score variance of each language, it is proposed that including normalisation coefficients for cross-language polarity comparison will help improve the approach	English, Arabic, Chinese, French, German, Italian, Japanese, Korean, Spanish	√	√	√	[42]
Lexicon and corpus-based with pattern transfer translation to identify a set of sentiment units	• Coverage of patterns is important for the accuracy of the approach • It is essential to understand the knowledge and techniques of text translation to derive parsing rules and patterns	English, Japanese	√	√	√	[45]
Concept-based with translation	• Disambiguation algorithms for identifying the context within text • Manual effort is needed as the polarity of some concepts may be 'opposite' in nature • Translation errors, untranslated terms and out-of-vocabulary (OOV) concepts	English, Chinese			√	[49]

Table 2. Lexicons and corpora used in multilingual sentiment analysis

Language	Name	Type	Remarks	Reference
English	OpinionFinder [5]	Subjectivity lexicon	6,856 unique entries, out of which 990 are multi-word expressions and with attributes – strong, weak and word senses (verb, adj, adv)	[3]
English	Negation terms Valenceshifters.tff [5], [22]	Negation lexicon	88 negation terms	[32]
English	244 intensifier Intensifiers2.tff [5], [22]	Intensifier lexicon	244 intensifiers	[32]

Language	Name	Type	Remarks	Reference
English	SentiWordNet [30]	Polarity lexicon	Trio of polarity scores assigned (positivity, negativity and objectivity scores); the sum of these scores is always 1	[9], [44]
English	Subjectivity clues Subjclueslen1-HLTEMNLP05.tff [5], [22]	Polarity lexicon	2718 English positive and 4910 negative terms	[32]
English	LingPipe movies reviews[4]	Polarity corpus	1000 positive and 1000 negative reviews	[9], [39], [41]
English	MPQA [31]	Polarity corpus	535 news articles from 187 different foreign and U.S. news sources; 4,958 sentences (1,471 positive and 3,487 negative)	[9], [33], [34]
English	Multi-domain sentiment corpus [50]	Polarity corpus	8,000 Amazon product reviews (4000 positive + 4000 negative)	[8], [41]
English	NTCIR Opinion Analysis Pilot Task [35], [36]	Polarity corpus	1,737 sentences (528 positive and 1,209 negative)	[33], [34]
English	NTCIR 8 Multilingual Opinion Analysis Task (MOAT)[5]	Polarity corpus	6,223 opinion units	[44]
English	General Inquirer Categories[6]	Polarity Dictionary		[24]
English	WordNet [51]	Vocabulary lexicons with synonym sets		[40]
English	Reuters RCV1[7]	Financial training corpus	800,000 texts, each containing 200-400 words	[21]
English	British National Corpus (BNC)[8]	General language		[21], [40]
Chinese	HIT IR-Lab Tongyici Cilin [52]	Lexical database	77,3443 Chinese words within 17,817 synsets	[53]
Chinese	HowNet Chinese sentiment lexicon[9]	Polarity lexicon	60,000 Chinese words and 11,000 sentences	[32], [41], [53]
Chinese	Product reviews[10]	Polarity corpus	886 IT product reviews (451 positive and 435 negative)	[8]
Chinese	NTCIR Opinion Analysis Pilot Task [36]	Polarity corpus	4,294 sentences (2,378 positive and 1,916 negative)	[33], [34]

[4] http://alias-i.com/lingpipe/index.html
[5] http://research.nii.ac.jp/ntcir/ntcir-ws8/permission/ntcir8xinhua-nyt-moat.html
[6] http://www.wjh.harvard.edu/~inquirer/homecat.htm
[7] http://trec.nist.gov/data/reuters/reuters.html
[8] http://www.natcorp.ox.ac.uk/corpus/
[9] http://www.keenage.com/
[10] http://www.it168.com/

Language	Name	Type	Remarks	Reference
Chinese	Douban reviews[11]	Polarity corpus	Movie/music/book reviews with 1000 positive and 1000 negative for each domain	[41]
Chinese	OPINMINE Chinese opinion annotation corpus [54]	Polarity corpus	Annotation corpus from NTCIR-6 Opinion Analysis Task	[53]
Chinese	Bing online English-Chinese dictionary[12]	Dictionary		[53]
Chinese	English-to-Chinese dictionary LDC_CE_DIC2.0 [32]	Dictionary	128,366 Chinese terms and their corresponding English terms	[32]
Chinese	Localisation for Taiwan and Big5 Encoding (TaBE)[13]	General language		[21]
German	Movie reviews[14]	Polarity corpus		[40]
English Swedish	The People's Dictionary[15]	English-Swedish dictionary		[16]
English Romanian	SemCor corpus [19]	Annotated subjective parallel corpus	Parallel corpus of 107 documents covering topics in sports, politics, fashion, education and others	[3]
English Romanian	Romanian NLP[16]	Various resources for Romanian NLP	Corpus of newspaper articles (50 million words), sense tagged data (39 ambiguous words), Romanian-English parallel text (1 million words), Romanian-English dictionary (38,000 entries)	[3]
Chinese-English	StarDict[17]	Dictionary	10 bilingual lexicons	[24]
English, German, French, Japanese	Cross-Lingual Sentiment (CLS) dataset[18]	Polarity corpus	800,000 Amazon product reviews in four languages (the product categories are books, dvds and music)	[37]
English, Chinese, German	Amherst Sentiment Corpus [55]	Polarity corpus		[40]
English-German	Ding[19]	Dictionary		[40]

[11] http://www.douban.com/
[12] http://cn.bing.com/dict/
[13] http://sourceforge.net/projects/libtabe/
[14] http://www.cs.colorado.edu/~jbg/static/data.html
[15] http://folkets-lexikon.csc.kth.se/folkets/folkets.en.html
[16] http://web.eecs.umich.edu/~mihalcea/downloads.html#romanian
[17] http://goldendict.org/dictionaries.php
[18] http://www.uni-weimar.de/en/media/chairs/webis/corpora/webis-cls-10/
[19] https://www-user.tu-chemnitz.de/~fri/ding/

Language	Name	Type	Remarks	Reference
Chinese-English	MDBG[20]	Dictionary		[40]
Chinese-German	HanDe[21]	Dictionary		[40]
Multiple	Universal Dictionary download site[22]	Dictionary	Web volunteer contributors – 4,500 entries in Romanian	[3]

3. Other work on multilingual analysis

The scope of multilingual analysis does not restrict to subjectivity and polarity analysis; it also includes cross-language document summarisation [56] and information retrieval in web search [57], among others. A list of relevant tools for multilingual sentiment analysis is shown in Table 3, as a reference for other cross-language studies. Briefly, two types of resources are shared, i.e., translators and unlabelled parallel corpora. Besides the commonly used translator such as Google Translate, commercial[23] and open source [58] tools are also covered. A number of studies have shown that Yahoo Babel Fish[24] contains the least 'correct' translation after manual inspection (e.g., see [32], [44]), and hence it is often used as a baseline either for manual correction or to impede translation bias of a human translator [44]. Parallel corpora can be a valuable asset for learning and overcoming cultural and linguistic diversity, so that information can be shared accurately and transparently across different societies with different languages.

Table 3. Tools for multilingual analysis

Type	Name	Language	Reference
Translator	PROMT eXcellent Translation (XT) Technology[25]	German, English, Spanish, French, Portuguese, Italian and Russian	[9]
Translator and Mapping	Google Translate[26]	Multiple Chinese-to-English tasks of MT2005, the BLEU-4 score~ is 0.3531 [32]	[8], [32], [53]
Translator	Yahoo Babel Fish[27]	Multiple Chinese-to-English tasks of MT2005, the BLEU-4 score~ is 0.1471 [32]	[32]
Translator	Bing Translator[28]	Multiple	[44]
Translator	Moses [58]	Multiple	[44]
Translator	IBM WebSphere Translation Server (WTS)[29]	Multiple	[42]

[20] http://www.mdbg.net/chindict/chindict.php
[21] http://www.handedict.de/
[22] http://www.dicts.info/uddl.php
[23] http://www.promt.com/
[24] http://www.babelfish.com/
[25] http://www.promt.com/
[26] https://translate.google.com.au/
[27] http://www.babelfish.com/
[28] http://www.bing.com/translator/

Type	Name	Language	Reference
Unlabelled Parallel Corpus	ISI Chinese-English Parallel Corpus [59]	English, Chinese	[33], [34]
Unlabelled Parallel Corpus	Parallel Corpora of 23 Official EU Languages[30]	Multiple	[42], [60]
Unlabelled Parallel Corpus	Parallel Corpora of 21 European Languages (extracted from the proceedings of the European Parliament) [61]	Multiple	[2], [40]

~BLEU (Bilingual Evaluation Understudy) is an algorithm for evaluating the quality of a piece of translated text. The score is calculated based on a modified form of precision for comparing the candidate translation against multiple references.

4. Sentiment analysis on social media

In the earlier days, companies and governments did not realise the power of social media, until they saw the influence of word-of-mouth and how quickly it could resonate with the community and inspire the launch of a protest or campaign for a cause [62]. Since then, sentiment analysis has expanded from being a research area on formal languages such as English to include informal languages used on social media. In particular, the content of tweets (posts shared on Twitter) is among the most studied, due to their ability to propagate hot topics in a very short duration and to a large number of users over wide geographical regions.

However, as seen from Section 2, most of the sentiment analysis studies to date have utilised resources such as lexicons and manually labelled corpora in English. The corpora used are mainly from news [3], [31], [33], [34] and reviews [8], [9], [37], [39], [41], [63], with content written in proper English. Given that social media is becoming the mainstream mode for communication and expressing one's thoughts on a variety of issues, it is essential to analyse the structure of social media corpora and current approaches used for sentiment analysis and opinion mining on social media.

4.1 English sentiment analysis on social media

Even though it is common for tweets to include many linguistic variations or mixed languages (especially in multicultural societies), most sentiment analysis studies still focus on English content because of the availability of resources. Pak and Paroubek [64] collected a corpus of 300,000 text posts from Twitter for objectivity and positive/negative emotion analysis. They concluded that Twitter users would use syntactic structures to describe emotion or state facts, and that Part-of-Speech (POS) tags may be strong indicators of emotional text. In addition, there is a difference in using the POS tags when expressing different types of emotion; positive text uses mostly superlative adverbs, such as "most", "best" and possessive endings, while negative text contains more verbs in

[29] http://www-03.ibm.com/software/products/en/translation-server
[30] https://ec.europa.eu/jrc/en/language-technologies/jrc-acquis

the past tense. An MNB classifier with n-grams and POS tags as features was tested and they found that the best performance is achieved with using bigrams.

Like Pak and Paroubek [64], Barbosa and Feng [65], Kouloumpis et al. [66] and Davidov et al. [67] followed the machine-learning based approach for Twitter sentiment analysis. Barbosa and Feng [65] proposed a two-step approach to classify the sentiment of tweets using SVM classifiers with abstract features. Kouloumpis et al. [66] evaluated training data extracted from hashtags and emoticons and examined if Twitter features play an important role in Twitter sentiment analysis. Davidov et al. [67] used a supervised k-nearest neighbours-like classifier to classify tweets into multiple sentiment types using hashtags and smileys as labels.

In contrast, Jiang et al. [68] classified the sentiment of a tweet according to its positive, negative or neutral sentiment about a target or entity. They argued that the context of a tweet is important to understand the underlying sentiment, and hence related tweets should be taken into consideration rather than just relying on a single tweet, which is usually too short and ambiguous for sentiment analysis. They used Pointwise Mutual Information (PMI) [69] to identify the extended target and implemented a three-level approach for detecting subjectivity, polarity and graph-based relationships. Their results show that the proposed approach is able to improve the performance of target-dependent sentiment classification.

4.2 Multilingual sentiment analysis on social media

While the aforementioned studies concentrate on sentiment analysis with English content, there are also research studies that use tweets as a corpus for multilingual sentiment analysis. Volkova et al. [70] proposed an approach for bootstrapping subjectivity clues from Twitter data and evaluated their approach on English, Spanish and Russian Twitter streams. The proposed approach uses the MPQA lexicon [22] to bootstrap sentiment lexicons from a large pool of unlabelled data using a small amount of labelled data to guide the process. Terms that are strongly subjective in translation are used as seed terms in the new language, with term polarity projected from the English lexicon. However, it is challenging to classify subjective tweets with philosophical thoughts. This is mainly due to some terms being weakly subjective and hence useable on both neutral and subjective tweets. Besides that, terms with ambiguous word sense and contradicting polarity (depending on the context) are found to be particularly error-prone.

Balahur and Turchi [15] built a simple sentiment analysis system for tweets in English, and used tweets from SemEval 2013 Task 2 – Sentiment Analysis in Twitter [71] as their training and testing datasets. They also translated the datasets from English to four other languages – Italian, Spanish, French and German. It is found that joint training datasets from languages with similar structures help to achieve improvement over the results obtained on an individual language. While this method is attractive, as it helps to disambiguate the contextual use of specific words, it cannot eliminate the error introduced by translation. From the findings, it is clear that considering the different ways the negation terms are constructed in the different languages is highly essential.

Cui et al. [72] did not use a translation machine but instead, focused on building emotion tokens or SentiLexicon using emoticons, repeating punctuations and repeating letters. These emotion tokens are first extracted to build a co-occurrence graph and through a graph propagation algorithm, positive and negative lexicons are labelled. The type of language is identified through Unicode of the character. If a tweet is from the Basic Latin or symbols section, it is assigned as a Basic Latin tweet. Most of the tweets considered are English in nature. Those characters in the Latin extended section are often in Portuguese, Spanish, German, and so on. Their comparative evaluation with SentiWordNet [30] indicated that emotion tokens are helpful for both English and non-English Twitter sentiment analysis.

4.3 Discussion

It is worth highlighting that approaches from sentiment analysis on social media are mainly based on pattern discovery such as syntactic structures [64], Twitter features [66], [67], emotion tokens [72], machine learning through annotated datasets [65], [67], [68] and translation [15], [70]. Considering the limited words available in a tweet and its evolving vocabulary, it is not surprising that the parallel corpus-based approach is not adopted as compared to other multilingual studies discussed in Section 2. In addition, sentence structures or grammatical rules are hardly considered even though Pak and Paroubek [64] showed that POS tags can be a useful indicator of emotional text. POS tags are only applicable if the subject of study is of a single language with proper grammatical rules, as the identification of tags is not straightforward when a tweet contains a mixture of languages. In fact, Kouloumpis et al. [66] showed that POS features may not be useful for sentiment analysis but other features such as emoticons and intensifiers are more useful in comparison. It is observed that none of the multilingual sentiment analysis studies takes into consideration the multiple languages found in a tweet. Instead, their focus is typically on studying the effects of different languages on a Twitter platform [70], [72] or leveraging available resources of one language for sentiment analysis of another language [15].

5. Work on scarce resource languages

In addition to the informal languages used on social media, as discussed in the previous section, this section explores studies that analyse languages with limited electronic resources, i.e., either no available or very minimal Natural Language Processing (NLP) tools can be found for the language. In a review paper such as this, it is important to consider and try to understand research that has been done on scarce resource languages. On top of developing NLP tools for some of those languages [73], efforts have also been made in the following three areas: sentiment analysis itself [74]-[77], speech recognition [78], [79] and machine translation [80], [81]. While studies on sentiment analysis along this line of research often concentrate on developing resources and approaches for a single scarce resource language, the other areas, speech recognition and machine translation, also look into constructing resources for other languages in order to support their research (such as crowd sourcing [80]).

5.1 Sentiment analysis

As in multilingual sentiment analysis, subjectivity analysis and polarity analysis have been done on scarce resource languages, although not extensively due to the limited resources available. Banea et al. [74] created a subjectivity lexicon for the Romanian language using a small set of seed words, a basic dictionary, and a small raw corpus. They used a bootstrapping approach to add new related words to a candidate list. They also used both PMI [82], [83] and Latent Semantic Analysis (LSA) [84] to filter noise from the lexicon. The caveat of their approach is that the LSA module needs to be trained using a sufficiently large corpus, and it is suggested that semi-automatic methods should be used for corpus construction as proposed by Ghani et al. [85]. Banea et al. showed that unsupervised learning using a rule-based sentence level subjectivity classifier is able to achieve a subjectivity F-measure score of 66.2, which is an improvement compared to previously proposed semi-supervised methods.

Bakliwal et al. [75] constructed a Hindi subjective lexicon for polarity classification of Hindi product reviews. Using WordNet [27] and a graph-based traversal method, they built a full (adjective and adverb) subjective lexicon. Their approach uses a small seed list with polarity to leverage the synonym and antonym relations of WordNet in order to expand on the initial lexicon. The subjectivity lexicon is then used in review classification. They achieved 79% accuracy using unigram and polarity scores as features. Another approach by Chowdhury and Chowdhury [76] uses both Bengali and English words to perform sentiment analysis on tweets. They applied a semi-supervised bootstrapping method to create the training corpus for machine learning classification, and achieved 93% accuracy through an SVM using unigrams with emoticons as features.

The study by Souza and Vieira [77] concentrated on sentiment analysis of Portuguese tweets using Portuguese polarity lexicons and negation models. They found that different lexicons such as Oplexicon and SentiLex actually have different accuracies. Specifically, Oplexicon [86] has better performance compared to SentiLex [87], due to the former's more comprehensive coverage of types of words and domains. A separate study by Elming et al. [88] used a robust offline-learning approach for cross-domain sentiment analysis on Danish based on a polarity lexicon. They observed significantly poorer performance when the analysis is done from one domain to another (i.e., reviews from the film domain to the company domain).

As shown above, the efforts in analysing sentiment on scarce resource languages are predominately devoted to constructing polarity lexicons [74], [75], [76] or making use of an available lexicon for sentiment classification [77], [88]. This is understandable as lexicon-based approaches are also widely adopted in multilingual sentiment analysis (see Section 2). Part of the reason being that, a polarity lexicon provides a straightforward method in assigning polarity to some content depending on the existence of a term or terms. This offers a viable option given the constraint of other resources, such as the availability of synonym dictionaries and translation machines.

5.2 Speech recognition

Thomas et al. [78] proposed to train deep neural networks (DNNs) [89] for low resource speech recognition. To overcome the limitation of having insufficient training data, they used transcribed data from other languages to build multilingual acoustic models. They observed a 16% improvement with just one hour of in-domain training, and three-fourths of the gain comes from DNN-based features.

Qian et al. [79] used a data borrowing strategy and the Subspace Gaussian Mixture Model [90] for the same problem. Even though their approach achieves only an improvement of about 1.7%, the results indicate that it is important to select languages that are linguistically similar and tie parameters at a context-dependent state.

5.3 Machine translation

Machine translation approaches often rely on parallel corpora to improve their accuracy and coverage. However, limited resources available for some of the languages imply that developing a machine translation engine can be an expensive task in terms of money and effort spent. Human annotation efforts and the availability of experts are required for the success of such tasks. Ambati et al. [80] proposed an approach to leverage active learning of 'sentence selection' through crowd-sourcing to enable automatic translation of low-resource language pairs. While the use of Mechanical Turk for annotation tasks has always been questioned, Ambati et al. showed that it is possible to create parallel corpora using non-experts with sufficient quality assurance.

In contrast, Irvine and Callison-Burch [81] used comparable corpora to improve the accuracy of translation from a small parallel corpus. They utilised a bilingual lexicon induction technique to learn new translation from the comparable corpora using a phrase-based statistical machine translation model for six low resource languages. Their results indicate that adding induced translation of low frequency words can improve the performance beyond inducing OOVs alone.

6. Challenges and recommendations

As shown in Table 1, common challenges encountered in multilingual sentiment analysis research include the word sense ambiguity problem [3], [9], [21], [49], language specific structure (negation [15] or parsing rules [45]) and translation errors [8], [9]. Most of the challenges are relevant to scarce resource languages, except for the errors introduced by translation machines, as most of these languages do not have such machines available to them.

6.1 Word sense dis-ambiguity

There are various suggestions for addressing the word sense ambiguity problem. Xia et al. [49] used Latent Dirichlet Allocation (LDA) [91] to extract top words that are related to a topic, and adopted PMI [82], [83] to calculate the polarity tendency of an opinion. Banea et al. [74] suggested that LSA [84] is sufficient to calculate the similarity between an original seed and each of the candidates extracted through a bootstrapping process. Active learning [80], which is used to improve machine translation by selecting sentences that are most informative for the task at hand, may help in targeting phrases or improving sample selection. These phrases and samples collected can be useful for a manual dis-ambiguity annotation process and also as input for feedback learning of a machine learning approach.

6.2 Language structure

It is well-known that different languages have their own unique ways of expression; for example, it is found that in the Russian language, philosophical thoughts and opinions are often misclassified and hence lexicon-based approaches may not be sufficient [70]. Instead, a deeper linguistic analysis is required. In addition, negation rules may be different for different languages and hence may cause unnecessary errors [15]. For scarce resource languages, some of the variants or dialects can be quite different in nature [92]. In view of the fact that there are a total of 48 variants of English available around the world[31], with some being a mixture of languages, and others being non-native pronunciation of English as well as a host of other permutations, it is essential to understand the

[31] http://en.wikipedia.org/wiki/List_of_dialects_of_the_English_language

structure of a language such as these in order to assess the best approach for leveraging the available English sentiment analysis resources.

6.3 Machine learning

Most of the scarce resource languages are used on social media, where slang or informal languages and emoticons are commonly found. A number of research studies have been able to achieve reasonably good results by including emotion tokens as features in their machine learning approaches [67], [72], [76]. Read [93] studied emoticons using text from the Usenet newsgroups. He classified the text into positive and negative types with both the SVM and NB, and achieved an accuracy of around 70% on the test set used. Go et al. [94] used a similar idea but they constructed their corpus from tweets. The best result of 81% accuracy was obtained using the NB classifier. These methods, however, do not perform well in identifying neutral text. A multi-level/cascading [39] or meta-classifier [44] approach has therefore been recommended for multilingual sentiment analysis where subjectivity analysis should be done before polarity analysis is conducted.

6.4 Essential resources

Subjectivity analysis cannot be accomplished without a lexicon or annotated corpus. Even though most of the scarce resource languages have limited resources available, an initial annotated dictionary or lexicon is still needed before a classifier with reasonable accuracy can be achieved. The following are two proposed approaches for creating lexicons for scarce resource languages, depending on the availability of resources:

1. A small bilingual dictionary as the available resource
   The only way to construct a subjective lexicon is by translating an existing lexicon from another language through the use of a bilingual dictionary. Although this mapping process can be automated, the accuracy would unfortunately be rather low due to the coverage limitation of the initial dictionary and the context-free translation process, which can introduce many word ambiguity problems. It is essential for the created lexicon to be verified by human annotators to ensure its quality, so that it can be used as a basis for generating more resources for a given scarce resource language.

2. A small subjective lexicon as the available resource
   A set of seed words can be selected from the lexicon to extract a corpus containing the seeds via a keyword search on the content of interest. From this set of candidates, a bootstrapping method can be applied, with their relatedness being measured using similarity metrics such as LSA or PMI to increase the volume of the lexicon. We recommend using the bootstrapping algorithm specified in the work of Banea et al. [74], if a reasonably-sized dictionary is available, or adopting the approach by Volkova et al. [70] to extract subjectivity lexicons from social media content, which is typically short and relatively non-structured.

The review from Section 4 indicates that none of the multilingual sentiment analysis studies on social media takes into account the possibility of having mixed languages in messages shared, even though it is common for social media data to have such languages (e.g., Singlish with words from English, Malay and Chinese dialects in a single tweet [92], [95]). It is therefore necessary to consider a more comprehensive polarity lexicon that contains polarity lexicons for each of the languages. As mentioned in Section 6.2, negation rules may be different for different languages. However, due to the extensive effort required for parsing a sentence in a scarce resource language, initiatives in identifying the different negation terms can be rewarding as a start. These negation terms can be coupled with the combined lexicon built for more accurate classification. Future work should investigate the behaviour and structure of sentences of different languages in order to construct a list of knowledge-based negation rules.

6.5 A hybrid framework

In view of the limitation of resources and challenges discussed, it is worth exploring a framework that incorporates both knowledge-based techniques (e.g., polarity lexicons) and statistical methods (e.g., machine learning) [96]. The recommended hybrid framework is shown in Figure 1. This proposed framework is especially applicable to scarce resource languages, when resources such as polarity lexicons and dictionaries may not be available or comprehensive enough. As can be seen from the figure, machine learning can be used for assigning polarity if that is the case. Even though it is a requirement to have an annotated training dataset before a machine learning model can be generated, semi-supervised methods with the use of emoticons (see Section 6.3 and references therein) or hashtags [66], [67] to extract a preliminary dataset with polarity can be adopted before manual annotation is done. The established hybrid framework is able to assign polarity to unseen content (considering the situation when none of the words matches any term in a polarity lexicon) by learning hidden rules of the annotated data. In addition, unseen data that has been classified can be reviewed for knowledge-based rule extraction or as a feedback system to improve machine learning classification.

Due to scarce resource limitations and multilingual settings, this framework can be adapted depending on resources available and the target language(s) to be analysed. The polarity pattern mentioned in Figure 1 can be a polarity word found in a lexicon or a type of negation pattern specific to a language. The knowledge-based polarity assignment is mainly based on resources or algorithms developed through detailed analysis of the language or languages. It can be a mixed language lexicon to address the mixture of languages found in social media data and/or knowledge-based negation rules mentioned in Section 6.4. In addition, the knowledge learnt from word sense dis-ambiguity explained in Section 6.1 can be incorporated into the knowledge-based algorithm to improve the accuracy of polarity assignment. The machine learning polarity assignment can adopt a simple model trained using a training dataset with emoticons or ensemble/cascading learning pointed out in Section 6.3. The accuracy of the proposed framework is heavily dependent on the final approaches implemented in the various components and quality of resources available.

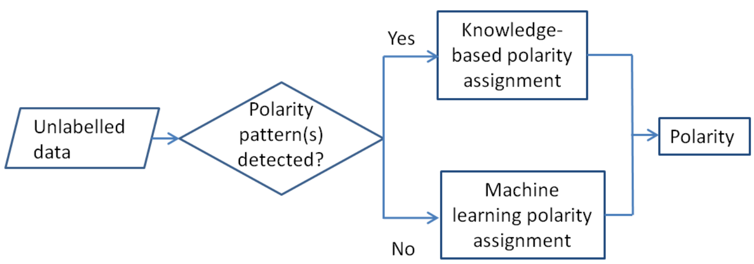

Figure 1. The recommended hybrid framework

6.6 Other considerations

While a manually annotated lexicon or corpus is still vital for sentiment analysis, it requires financial funding support and a considerable amount of human effort to create a reasonably sized resource. If funding is not a limiting factor, a crowd-sourcing approach [80], [97] can be considered, as the quality of annotation can be improved through cross validation and verification of several annotators. However, if crowd sourcing is not a viable option, an initial polarity corpus can be created by using emotion tokens [72]. This corpus can then be put together with the lexicon built, to discover more candidates through a bootstrapping [74] or SCL [37] approach. It is common to use a subjectivity lexicon for a rule-based classifier, however, a number of studies [2], [42] have shown that a combination of corpus-based machine learning and lexicon rule-based methods with cascading learning [39] can improve the accuracy of sentiment analysis.

Even though the linguistic structure of a scarce resource language is important for determining if English resources can be adapted successfully, it requires detailed analysis to be carried out by linguistic experts in order to identify the structural differences. As a result, it is suggested that machine learning should be used as an alternative or a litmus test, to assess if there is a need for a structural study to further improve the accuracy. As shown in Table 1, one of the downfalls is the limited ability of a classifier to recognise negative text and omitted negation structures. While it may not be possible to conduct a study on the linguistic structure of a scarce resource language, it is certainly possible to manually identify some negation samples from the available corpus and incorporate the specific pattern or structure when constructing a training dataset.

To sum up, although the lexicon-based approach is still essential for sentiment analysis, it should be expanded to include contextual awareness features, as most of the sentiments are related to an entity or a topic. Apart from that, the concept-based approach, which incorporates common-sense reasoning [98], is fast developing. Concept-level sentiment analysis is necessary for managing more subtle sentiments that are often not captured or handled in current multilingual sentiment analysis research.

7. Conclusion

Sentiment analysis is an active research area, thanks to the many challenges but also promises. While many sentiment analysis studies have been conducted on formal languages using mainstream platforms like news or official documents, increasing attention is now placed on analysis of social media content to facilitate understanding of the wellbeing of a community or the perceived image of a company/product. Social media content often contains informal or mixed languages. It is thus no longer sufficient to consider only a formal language (e.g., English) in sentiment analysis research. In this review paper, we have looked at a range of current approaches and tools used for multilingual sentiment analysis. We took into account not just formal languages but informal and scarce resource languages too. Major challenges have been identified, and we recommended possible remedies as well as a hybrid framework for developing sentiment analysis resources particularly for languages with limited electronic resources.

References

[1] E. Riloff and J. Wiebe, 'Learning extraction patterns for subjective expressions', in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2003, pp. 105–112.

[2] S.-M. Kim and E. Hovy, 'Identifying and analyzing judgment opinions', in *Proceedings of the conference of North American Chapter of the Association of Computational Linguistics*, 2006, pp. 200–207.

[3] R. Mihalcea, C. Banea, and J. Wiebe, 'Learning multilingual subjective language via cross-lingual projections', in *Proceedings of Annual Meeting of Association for Computational Linguistics*, 2007, vol. 45, p. 976.

[4] B. Pang and L. Lee, 'Opinion mining and sentiment analysis', *Found. Trends Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008.

[5] T. Wilson, P. Hoffmann, S. Somasundaran, J. Kessler, J. Wiebe, Y. Choi, C. Cardie, E. Riloff, and S. Patwardhan, 'OpinionFinder: A system for subjectivity analysis', in *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 2005, pp. 34–35.

[6] H. Kanayama and T. Nasukawa, 'Fully automatic lexicon expansion for domain-oriented sentiment analysis', in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2006, pp. 355–363.

[7] Y. Hu, J. Duan, X. Chen, B. Pei, and R. Lu, 'A new method for sentiment classification in text retrieval', in *Proceedings of International Joint Conference on Natural Language Processing*, 2005, pp. 1–9.

[8] X. Wan, 'Co-training for cross-lingual sentiment classification', in *Proceedings of the Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing*, 2009, pp. 235-243.

[9] K. Denecke, 'Using sentiwordnet for multilingual sentiment analysis', in *Proceedings of International Conference on Data Engineering Workshops*, 2008, pp. 507–512.

[10] J. R. Leimgruber, 'Singapore English', *Lang. Linguist. Compass*, vol. 5, no. 1, pp. 47–62, 2011.

[11] E. Cambria, D. Olsher, and D. Rajagopal, 'SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis', in *Proceedings of AAAI Conference on Artificial Intelligence*, 2014, pp. 1515–1521.

[12] S. Tan and J. Zhang, 'An empirical study of sentiment analysis for chinese documents', *Expert Syst. Appl.*, vol. 34, no. 4, pp. 2622–2629, 2008.

[13] J. Zhao, L. Dong, J. Wu, and K. Xu, 'Moodlens: an emoticon-based sentiment analysis system for chinese tweets', in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012, pp. 1528–1531.

[14] N. Kobayashi, K. Inui, Y. Matsumoto, K. Tateishi, and T. Fukushima, 'Collecting evaluative expressions for opinion extraction', in *Proceedings of International Conference on Natural Language Processing*, 2005, pp. 596–605.

[15] A. Balahur and M. Turchi, 'Improving sentiment analysis in Twitter using multilingual machine translated data.', in *Proceedings of Recent Advances in Natural Language Processing*, 2013, pp. 49–55.

[16] M. Rosell and V. Kann, 'Constructing a swedish general purpose polarity lexicon random walks in the people's dictionary of synonyms', in *Proceedings of Swedish Language Technology Conference*, 2010, pp. 19–20.

[17] M. Abdul-Mageed, M. T. Diab, and M. Korayem, 'Subjectivity and sentiment analysis of modern standard arabic', in *Proceedings of the Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers*, 2011, vol. 2, pp. 587–591.

[18] E. Cambria and A. Hussain, *Sentic computing: a common-sense-based framework for concept-level sentiment analysis*, vol. 1. Springer, 2015.

[19] G. A. Miller, C. Leacock, R. Tengi, and R. T. Bunker, ‘A semantic concordance’, in *Proceedings of the Workshop on Human Language Technology*, 1993, pp. 303–308.
[20] D. D. Lewis, ‘Naive (Bayes) at forty: The independence assumption in information retrieval’, in *Proceedings of European Conference on Machine Learning*, 1998, pp. 4–15.
[21] K. Ahmad, D. Cheng, and Y. Almas, ‘Multi-lingual sentiment analysis of financial news streams’, in *Proceedings of the International Conference on Grid in Finance*, 2006.
[22] T. Wilson, J. Wiebe, and P. Hoffmann, ‘Recognizing contextual polarity in phrase-level sentiment analysis’, in *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 2005, pp. 347–354.
[23] A. Esuli and F. Sebastiani, ‘Determining Term Subjectivity and Term Orientation for Opinion Mining.’, in *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics*, 2006, vol. 6, p. 2006.
[24] J. Yao, G. Wu, J. Liu, and Y. Zheng, ‘Using bilingual lexicon to judge sentiment orientation of Chinese words’, in *Proceedings of IEEE International Conference on Computer and Information Technology*, 2006, pp. 38–38.
[25] V. Vapnik, *The nature of statistical learning theory*. Springer Science & Business Media, 2000.
[26] J. R. Quinlan, *C4. 5: programs for machine learning*. Elsevier, 2014.
[27] G. A. Miller, ‘WordNet: a lexical database for English’, *Commun. ACM*, vol. 38, no. 11, pp. 39–41, 1995.
[28] F. J. Och and H. Ney, ‘Improved statistical alignment models’, in *Proceedings of the Annual Meeting on Association for Computational Linguistics*, 2000, pp. 440–447.
[29] V. Kann and M. Rosell, ‘Free construction of a free Swedish dictionary of synonyms’, in *Proceedings of the Nordic Conference on Computational Linguistics*, 2005, pp. 105–110.
[30] S. Baccianella, A. Esuli, and F. Sebastiani, ‘SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining.’, in *Proceedings of Language Resources and Evaluation Conference*, 2010, vol. 10, pp. 2200–2204.
[31] J. Wiebe, T. Wilson, and C. Cardie, ‘Annotating expressions of opinions and emotions in language’, *Lang. Resour. Eval.*, vol. 39, no. 2–3, pp. 165–210, 2005.
[32] X. Wan, ‘Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis’, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2008, pp. 553–561.
[33] X. Meng, F. Wei, X. Liu, M. Zhou, G. Xu, and H. Wang, ‘Cross-lingual mixture model for sentiment classification’, in *Proceedings of the Annual Meeting of the Association for Computational Linguistics: Long Papers*, 2012, vol. 1, pp. 572–581.
[34] B. Lu, C. Tan, C. Cardie, and B. K. Tsou, ‘Joint bilingual sentiment classification with unlabeled parallel corpora’, in *Proceedings of the Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 2011, vol. 1, pp. 320–330.
[35] Y. Seki, D. K. Evans, L.-W. Ku, H.-H. Chen, N. Kando, and C.-Y. Lin, ‘Overview of opinion analysis pilot task at NTCIR-6’, in *Proceedings of NTCIR-6 Workshop Meeting*, 2007, pp. 265–278.
[36] Y. Seki, D. K. Evans, L.-W. Ku, L. Sun, H.-H. Chen, N. Kando, and C.-Y. Lin, ‘Overview of multilingual opinion analysis task at NTCIR-7’, in *Proceedings of NTCIR-7 Workshop Meeting*, 2008.
[37] P. Prettenhofer and B. Stein, ‘Cross-lingual adaptation using structural correspondence learning’, *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 1, p. 13, 2011.
[38] J. Blitzer, R. McDonald, and F. Pereira, ‘Domain adaptation with structural correspondence learning’, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2006, pp. 120–128.
[39] E. Boiy and M.-F. Moens, ‘A machine learning approach to sentiment analysis in multilingual Web texts’, *Inf. Retr.*, vol. 12, no. 5, pp. 526–558, 2009.

[40] J. Boyd-Graber and P. Resnik, 'Holistic sentiment analysis across languages: Multilingual supervised latent Dirichlet allocation', in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2010, pp. 45–55.

[41] J. Pan, G.-R. Xue, Y. Yu, and Y. Wang, 'Cross-lingual sentiment classification via bi-view non-negative matrix tri-factorization', in *Advances in knowledge discovery and data mining*, Springer, 2011, pp. 289–300.

[42] M. Bautin, L. Vijayarenu, and S. Skiena, 'International sentiment analysis for news and blogs.', in *Proceedings of International Conference on Web and Social Media*, 2008.

[43] N. Godbole, M. Srinivasaiah, and S. Skiena, 'Large-scale sentiment analysis for news and blogs.', in *Proceedings of International Conference on Web and Social Media*, 2007, vol. 7, p. 21.

[44] A. Balahur and M. Turchi, 'Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis', *Comput. Speech Lang.*, vol. 28, no. 1, pp. 56–75, 2014.

[45] K. Hiroshi, N. Tetsuya, and W. Hideo, 'Deeper sentiment analysis using machine translation technology', in *Proceedings of the International Conference on Computational Linguistics*, 2004, p. 494.

[46] S. Poria, E. Cambria, G. Winterstein, and G.-B. Huang, 'Sentic patterns: Dependency-based rules for concept-level sentiment analysis', *Knowl.-Based Syst.*, vol. 69, pp. 45–63, 2014.

[47] E. Cambria, P. Gastaldo, F. Bisio, and R. Zunino, 'An ELM-based model for affective analogical reasoning', *Neurocomputing*, vol. 149, pp. 443–455, 2015.

[48] S. Poria, E. Cambria, A. Gelbukh, F. Bisio, and A. Hussain, 'Sentiment data flow analysis by means of dynamic linguistic patterns', *Comput. Intell. Mag. IEEE*, vol. 10, no. 4, pp. 26–36, 2015.

[49] Y. Xia, X. Li, E. Cambria, and A. Hussain, 'A localization toolkit for SenticNet', in *Proceedings of IEEE International Conference on Data Mining Workshops*, 2014, pp. 403–408.

[50] J. Blitzer, M. Dredze, and F. Pereira, 'Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification', in *Proceedings of Annual Meeting of Association for Computational Linguistics*, 2007, vol. 7, pp. 440–447.

[51] G. A. Miller, 'Nouns in WordNet: a lexical inheritance system', *Int. J. Lexicogr.*, vol. 3, no. 4, pp. 245–264, 1990.

[52] W. Che, Z. Li, and T. Liu, 'Ltp: A chinese language technology platform', in *Proceedings of the International Conference on Computational Linguistics: Demonstrations*, 2010, pp. 13–16.

[53] 'NTCIR8 MOAT Xinhua and NYT News corpus'. [Online]. Available: http://research.nii.ac.jp/ntcir/ntcir-ws8/permission/ntcir8xinhua-nyt-moat.html. [Accessed: 27-Mar-2015].

[54] R. Xu, K.-F. Wong, and Y. Xia, 'Opinmine–opinion analysis system by CUHK for NTCIR-6 pilot task', in *Proceedings of the NTCIR-6 Workshop*, 2007.

[55] N. Constant, C. Davis, C. Potts, and F. Schwarz, 'The pragmatics of expressive content: Evidence from large corpora', *Sprache Datenverarb.*, vol. 33, no. 1–2, pp. 5–21, 2009.

[56] F. Boudin, S. Huet, J.-M. Torres-Moreno, and J. Torres-Moreno, 'A graph-based approach to cross-language multi-document summarization', *Res. J. Comput. Sci. Comput. Eng. Appl. Polibits*, vol. 43, pp. 113–118, 2010.

[57] J. Savoy and L. Dolamic, 'How effective is Google's translation service in search?', *Commun. ACM*, vol. 52, no. 10, pp. 139–143, 2009.

[58] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, and R. Zens, 'Moses: Open source toolkit for statistical machine translation', in *Proceedings of the Annual Meeting on Association for Computational Linguistics : Demonstrations*, 2007, pp. 177–180.

[59] D. S. Munteanu and D. Marcu, 'Improving machine translation performance by exploiting non-parallel corpora', *Comput. Linguist.*, vol. 31, no. 4, pp. 477–504, 2005.

[60] 'IBM - WebSphere Translation Server for Multiplatforms'. [Online]. Available: http://www-03.ibm.com/software/products/en/translation-server. [Accessed: 28-Mar-2015].
[61] P. Koehn, 'Europarl: A parallel corpus for statistical machine translation', in *Proceedings of Machine Translation Summit*, 2005, vol. 5, pp. 79–86.
[62] W. Zhang, T. J. Johnson, T. Seltzer, and S. L. Bichard, 'The revolution will be networked: The influence of social networking sites on political attitudes and behavior', *Soc. Sci. Comput. Rev.*, 2009.
[63] 'LingPipe Home'. [Online]. Available: http://alias-i.com/lingpipe/index.html. [Accessed: 25-Mar-2015].
[64] A. Pak and P. Paroubek, 'Twitter as a corpus for sentiment analysis and opinion mining.', in *Proceedings of Language Resources and Evaluation Conference*, 2010, vol. 10, pp. 1320–1326.
[65] L. Barbosa and J. Feng, 'Robust sentiment detection on twitter from biased and noisy data', in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, 2010, pp. 36–44.
[66] E. Kouloumpis, T. Wilson, and J. D. Moore, 'Twitter sentiment analysis: The good the bad and the omg!', in *Proceedings of International Conference on Web and Social Media*, 2011, vol. 11, pp. 538–541.
[67] D. Davidov, O. Tsur, and A. Rappoport, 'Enhanced sentiment learning using twitter hashtags and smileys', in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, 2010, pp. 241–249.
[68] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, 'Target-dependent twitter sentiment classification', in *Proceedings of the Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 2011, vol. 1, pp. 151–160.
[69] Q. Su, K. Xiang, H. Wang, B. Sun, and S. Yu, 'Using pointwise mutual information to identify implicit features in customer reviews', in *Computer Processing of Oriental Languages. Beyond the Orient: The Research Challenges Ahead*, Springer, 2006, pp. 22–30.
[70] S. Volkova, T. Wilson, and D. Yarowsky, 'Exploring sentiment in social media: Bootstrapping subjectivity clues from multilingual twitter streams.', in *Proceedings of Annual Meeting of the Association of Computational Linguistics*, 2013, pp. 505–510.
[71] P. Nakov, Z. Kozareva, A. Ritter, S. Rosenthal, V. Stoyanov, and T. Wilson, 'Semeval-2013 task 2: Sentiment analysis in twitter', in *Proceedings of the International Workshop on Semantic Evaluation*, 2013.
[72] A. Cui, M. Zhang, Y. Liu, and S. Ma, 'Emotion tokens: Bridging the gap among multilingual twitter sentiment analysis', in *Information retrieval technology*, Springer, 2011, pp. 238–249.
[73] C. Monson, A. F. Llitjós, R. Aranovich, L. Levin, R. Brown, E. Peterson, J. Carbonell, and A. Lavie, 'Building NLP systems for two resource-scarce indigenous languages: Mapudungun and Quechua', *Strateg. Dev. Mach. Transl. Minor. Lang.*, p. 15, 2006.
[74] C. Banea, R. Mihalcea, and J. Wiebe, 'A bootstrapping method for building subjectivity lexicons for languages with scarce resources.', in *Proceedings of Language Resources and Evaluation Conference*, 2008, vol. 8, pp. 2–764.
[75] A. Bakliwal, P. Arora, and V. Varma, 'Hindi subjective lexicon: A lexical resource for Hindi polarity classification', in *Proceedings of Language Resources and Evaluation Conference*, 2012, pp. 1189–1196.
[76] S. Chowdhury and W. Chowdhury, 'Performing sentiment analysis in Bangla microblog posts', in *Proceedings of International Conference on Informatics, Electronics & Vision*, 2014, pp. 1–6.
[77] M. Souza and R. Vieira, 'Sentiment analysis on twitter data for portuguese language', in *Computational Processing of the Portuguese Language*, Springer, 2012, pp. 241–247.
[78] S. Thomas, M. L. Seltzer, K. Church, and H. Hermansky, 'Deep neural network features and semi-supervised training for low resource speech recognition', in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 6704–6708.

[79] Y. Qian, D. Povey, and J. Liu, 'State-level data borrowing for low-resource speech recognition based on subspace GMMs.', in *Proceedings of Annual Conference of the International Speech Communication Association*, 2011, pp. 553–560.

[80] V. Ambati, S. Vogel, and J. G. Carbonell, 'Active learning and crowd-sourcing for machine translation.', in *Proceedings of Language Resources and Evaluation Conference*, 2010, vol. 1, p. 2.

[81] A. Irvine and C. Callison-Burch, 'Combining bilingual and comparable corpora for low resource machine translation', in *Proceedings of the Eighth Workshop on Statistical Machine Translation*, 2013, pp. 262–270.

[82] P. D. Turney, 'Mining the Web for synonyms: PMI-IR versus LSA on TOEFL', *Lect. Notes Comput. Sci.*, pp. 491–502, 2001.

[83] P. D. Turney, 'Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews', in *Proceedings of Annual Meeting of the Association of Computational Linguistics*, 2002, pp. 417–424.

[84] S. T. Dumais, G. W. Furnas, T. K. Landauer, S. Deerwester, and R. Harshman, 'Using latent semantic analysis to improve access to textual information', in *Proceedings of the Special Interest Group on Computer-Human Interaction conference*, 1988, pp. 281–285.

[85] R. Ghani, R. Jones, and D. Mladenić, 'Mining the web to create minority language corpora', in *Proceedings of the International Conference on Information and Knowledge Management*, 2001, pp. 279–286.

[86] M. Souza, R. Vieira, D. Busetti, R. Chishman, and I. M. Alves, 'Construction of a portuguese opinion lexicon from multiple resources', in *Proceedings of the Brazilian Symposium in Information and Human Language Technology*, 2011, pp. 59–66.

[87] M. J. Silva, P. Carvalho, C. Costa, and L. Sarmento, 'Automatic expansion of a social judgment lexicon for sentiment analysis', 2010.

[88] J. Elming, D. Hovy, and B. Plank, 'Robust cross-domain sentiment analysis for low-resource languages', in *Proceedings of Annual Meeting of Association for Computational Linguistics*, 2014, pp. 2–7.

[89] L. Deng, G. Hinton, and B. Kingsbury, 'New types of deep neural network learning for speech recognition and related applications: An overview', in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 8599–8603.

[90] D. Povey, L. Burget, M. Agarwal, P. Akyazi, F. Kai, A. Ghoshal, O. Glembek, N. Goel, M. Karafiát, and A. Rastrow, 'The subspace Gaussian mixture model—A structured model for speech recognition', *Comput. Speech Lang.*, vol. 25, no. 2, pp. 404–439, 2011.

[91] D. M. Blei, A. Y. Ng, and M. I. Jordan, 'Latent dirichlet allocation', *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003.

[92] S. L. Lo, E. Cambria, R. Chiong, and D. Cornforth, 'A multilingual semi-supervised approach in deriving Singlish sentic patterns for polarity detection', *Knowl.-Based Syst.*, 2016.

[93] J. Read, 'Using emoticons to reduce dependency in machine learning techniques for sentiment classification', in *Proceedings of the Association for Computational Linguistics Student Research Workshop*, 2005, pp. 43–48.

[94] A. Go, R. Bhayani, and L. Huang, 'Twitter sentiment classification using distant supervision', *CS224N Proj. Rep. Stanf.*, pp. 1–12, 2009.

[95] S. L. Lo, R. Chiong, D. Cornforth, and Y. Bao, 'An unsupervised multilingual approach for identifying high-value topics on Twitter', Working Paper, 2016.

[96] E. Cambria, 'Affective computing and sentiment analysis', *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 102–107, 2016.

[97] E. Cambria, D. Rajagopal, K. Kwok, and J. Sepulveda, 'GECKA: game engine for commonsense knowledge acquisition', in *Proceedings of AAAI FLAIRS Conference*, 2015, pp. 282–287.

[98] E. Cambria, J. Fu, F. Bisio, and S. Poria, ‘AffectiveSpace 2: Enabling affective intuition for concept-level sentiment analysis’, in *Proceedings of AAAI Conference on Artificial Intelligence*, 2015, pp. 508–514.