

A Review on Multi-Lingual Sentiment Analysis by Machine Learning Methods

Santwana Sagnika^{1,*}, Anshuman Pattanaik¹, Bhabani Shankar Prasad Mishra¹ and Saroj K. Meher²

¹ School of Computer Engineering, Kalinga Institute of Industrial Technology Deemed to be University, Bhubaneswar, Odisha, India

² Systems Science And Informatics Unit, Indian Statistical Institute, Bangalore, Karnataka, India

Received 19 August 2019; Accepted 12 February 2020

Abstract

The arrival of e-commerce and the multitude of information presented by the web have established the internet as a principal destination for consumers looking for truthful opinions and multiple viewpoints for some product, news, topic, or trend in the markets. Thus, it is desirable to make this search easier by using systems which sift through the mass of data and summarize the available opinions for easy understanding of the seeker. This task, known as sentiment analysis, is currently a prominent area of research. Sentiment analysis can be useful for businesses, data analysts and data scientists, as well as customers. Even though many methods are designed to perform this task on English data, there is a lack of systems that can analyze data in other languages. This paper attempts to provide a detailed study on the sentiment analysis methods applied on languages other than English. The tools used, pros and cons, and efficiency of all methods is covered. The associated challenges are also discussed. The paper covers methods that analyze translated data as well as methods that analyze available data in the target language.

Keywords: Sentiment Analysis, Machine Learning, Deep Learning, Multi-Lingual, Cross-Lingual, Natural Language Processing, Web Data Mining, Text Mining

1. Introduction

With the increasing popularity of online platforms and social media in our daily lives, there is an information boom all over the internet. Hence, there is ample data on every topic, ranging from products, businesses, market trends, etc. Correspondingly, the opinions of the users of respective domains are also freely available in plenty. Be it movie reviews, product reviews, financial market sentiments, political opinions, etc., there is easy access to such data for anyone seeking to take an informed decision. In such a context, it becomes necessary as well as useful to have a mechanism that sifts through the mass of data, analyzes them and preferably categorizes, quantifies or scores them, in order to aid decision making. This has given rise to the concept of sentiment analysis, also known as opinion mining. Sentiment analysis is the process of analyzing opinions or views expressed in documents and their overall classification, scoring or quantification. The primary purpose is to get an idea of people's general attitude and feelings towards a certain subject [1, 2].

Sentiment analysis mostly deals with a huge amount of unstructured and unlabelled data. Besides, the available data is generally subjective, vague, and not strictly adherent to language rules. Hence, sentiment analysis becomes a complex task, and involves knowledge of various domains, including natural language processing, data mining, machine learning, data analytics, computational intelligence, etc. At the very basic level, sentiment analysis is a classification problem, which categorizes opinions into broad categories like positive, negative or neutral. But in-depth analysis can lead to

exploring finer details and extracting a larger amount of useful information [2, 3].

A lot of businesses need user feedbacks for decision making. They collect it using opinion polls, customer surveys, questionnaires, etc. With the widespread availability of the internet, such feedback collection methods are able to reach a broader range of consumers, who are able to provide their honest and unbiased opinions. The entire process is easier and takes minimal time. Hence, it helps businesses, service providers, e-commerce organizations, governments, etc. collect varied opinions and use them to aid decision-making processes [4].

Owing to the huge amount of text available online that expresses the views of users on common forums, manual processing and analysis of the text is cumbersome and time consuming. Automated Sentiment Analysis techniques help save manpower, get faster output, sift through massive unnecessary data to find relevant material, and present the results in necessary formats. Sentiment Analysis tools are greatly useful for extracting data from various sources, viz. review sites, feedback forums, social networking sites, blogs, and so on, and performing detailed analytical operations on it [4, 5].

Even though research on sentiment analysis has seen a lot of milestones, most of the work is performed on English data. In comparison, very less research has been performed on languages other than English. It is more critical to analyze non-English data and there are multiple challenges to this task. Mechanisms that work on other languages either rely on their own limited resources, or prefer translation to English and using the abundantly available English resources. This task, known as multi-lingual sentiment analysis, is still an open area of research with a considerable scope for improvement.

*E-mail address: santwana.sagnika@gmail.com

ISSN: 1791-2377 © 2020 School of Science, IITU. All rights reserved.

doi:10.25103/jestr.132.19

In this work, the authors review the available literature on sentiment analysis performed on languages other than English. A detailed survey is provided, along with the tools, techniques, mechanisms and performances, with special focus on machine learning techniques as they are highly popular in this field. The purpose of this is to provide a reader an introductory idea of the extent of research and analysis in the field, and provide scope for further research, as most of the sentiment analysis work is done in English, and rest of the languages are still open to research.

The paper is designed as follows. Section 2 elaborates on the basics of sentiment analysis. Section 3 gives an idea of the techniques employed. Section 4 deals with multi-lingual and cross-lingual analysis and the body of research available in this field. The focus is mostly on machine learning techniques. Section 5 concludes the work and suggests future work possibilities.

2. Preliminary Concepts

Sentiment Analysis is all about traversing through a substantial amount of unstructured data to find opinions and feelings expressed in them towards a certain topic or object. Sentiment Analysis can be considered as a sub-field of text processing.

Any text that comes under analysis can be either factual, i.e. expressing facts and information, or opinionated, i.e. expressing views and opinions. It is necessary to identify such opinionated statements, as they are relevant for opinion mining tasks. This process is known as Subjectivity Classification, which is a prerequisite of Sentiment Analysis. A subjective statement is that which expresses an emotion or feeling, which may or may not be an opinion, though most of the opinionated statements are generally subjective. Sometimes, even objective statements can carry opinions. Once an opinionated statement is discovered, Sentiment Analysis is performed on it, in order to classify the opinion as positive or negative. A deeper study is required to find out the type of opinion expressed in the statement. Hence, subjectivity classification and sentiment classification are complementary processes.

Broadly, Sentiment Analysis can be done at two levels. Document-level Sentiment Classification refers to finding the overall viewpoint presented in the document, whether positive or negative. Sentence-level Sentiment Classification is the process of analyzing individual sentences and finding their polarity, i.e. either positive or negative, after the sentences have been identified as opinionated ones [26]. It is noteworthy to state that a document can be considered as a collection of sentences, which may or may not express opinions. This means that a document may contain positive, negative and neutral statements. The individual analysis of these statements combines to find the overall polarity of the document [57].

A more recent approach considers Aspect-level Sentiment Classification, wherein individual topics are considered, and the sentiment regarding those particular topics is ascertained [36].

2.1. Tasks under Sentiment Analysis

Identifying and understanding the opinion expressed in texts is a highly complex task. There are many aspects involved. The general format of a sentiment analysis task specifies that a document can be expressed in terms of five components –

Opinion, Features, Object, Opinion Holder and Time of Expression.

- (i). Firstly it is imperative to find the target of the opinion, i.e. the object on which the opinion has been expressed. The opinion can be expressed on the object as a whole, or a specific component of an object, generally referred to as feature or aspect. A document can contain views towards multiple features of the same object, the object as a whole, and multiple objects as well. Some documents can contain comparative analysis of objects, based on some common features.
- (ii). Next, the type of opinion has to be detected. This is known as polarity identification. This is basically a classification task, categorizing opinions into broadly positive and negative. Finer-level classification can also help identify details, like the specific emotion associated with the opinion, e.g. joy, anger, sadness, surprise, etc.
- (iii). Some situations require finding the opinion holder, i.e. the person who expresses the opinion. It is generally assumed that a document contains the opinion of a single opinion holder throughout. But many times, there might be multiple people expressing their opinions, which may or may not match.
- (iv). Additionally, the time at which the opinion has been expressed becomes necessary in certain cases.
- (v). Identification of features influence the accuracy of the detection of polarity.

Features are the most significant factors while looking for sentiments. Some texts contain features visibly highlighted. For example, “The picture clarity of this TV set is excellent” clearly mentions that the view is being expressed about the “clarity” feature of the object “TV set”. These features are known as explicit features. Sometimes, implicit features are indicated, e.g. “The laptop is really heavy”. In this case, it has to be understood that the feature “weight” is being talked about, even if it isn’t mentioned anywhere. It is harder to analyze such texts.

The most important step in Sentiment Analysis, as is evident, is finding the polarity of an expressed opinion. For this purpose, the words are identified as “positive” or “negative”. Positive words include “good”, “excellent”, “satisfactory”, etc. Negative terms are like “poor”, “disgusting”, “bad”, etc. The presence of such terms gives an idea about the polarity of a statement. But this is not always such a straightforward task. The opinions can be expressed in many forms. Sometimes it is a clearly written *direct opinion*. For example, “The design of the phone is beautiful.” The direct opinion can also be expressed as an effect of an object on another, instead of using specific descriptive words, e.g. “This lamp brightens up the entire room”. Here, although not mentioned as good or bad explicitly, the sentence leads the reader to understand that the lamp serves its purpose, and is hence positively expressed by the user. Alternatively, *comparative opinion* can be expressed by the user, by performing a comparative analysis of two or more objects based on certain common features, and mentioning a preference based on the comparison, instead of dealing with a single object. For example, “Apple provides a better camera than Samsung.” Here, “better” is applicable on Apple as the object and not Samsung. Figure 1 shows the broad steps involved in a sentiment analysis task.

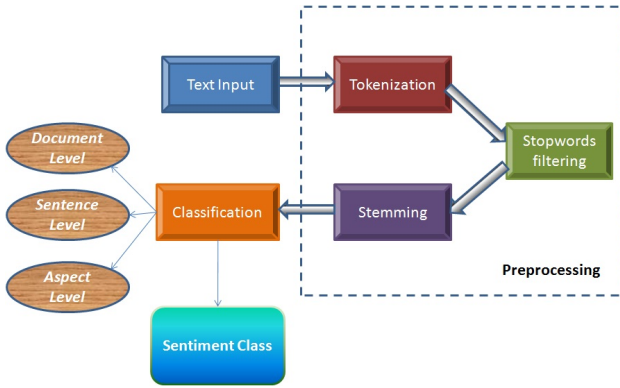


Fig. 1. Steps of Sentiment Analysis

2.2. Sources of text

The text on which sentiment analysis can be performed can be from a myriad of sources. Some of the most common sources are:

- (i). Opinion blogs – Blogging is a popular activity in today’s world. Blogs are available about a wide variety of topics like gadgets, current affairs, political issues, travel spots, etc. These can be analyzed to obtain opinions about the respective content.
- (ii). E-commerce reviews – Many users on popular e-commerce sites like Amazon, Flipkart, eBay, etc. give their feedback and reviews after using a certain product, generally accompanied with a star rating. These opinions can be referred by a prospective user who is looking for options to buy a product.
- (iii). Social media content and Media sharing – With approximately 2.5 billion users active on social media all over the world, it is obvious that the data generated on Facebook, Twitter, LinkedIn, Instagram, YouTube, SoundCloud etc. is rich and voluminous, and a likely place to look for information related to users’ opinions and choices.
- (iv). Communication – Communication mediums like SMS, WhatsApp etc. are also a rich source of information, which can be mined for information and opinions [48, 49].

3. Techniques for Sentiment Analysis

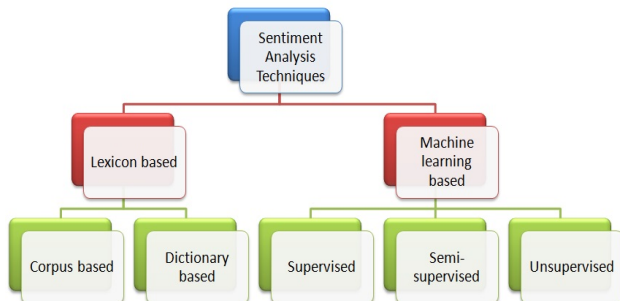


Fig. 2. Sentiment Analysis Techniques

Over the years, Sentiment Analysis has emerged as an interesting research problem, with multiple approaches being designed to solve sentiment analysis problems. These techniques can be broadly categorized into two major categories, as shown in figure 2 [7, 8, 9].

3.1. Methodologies

(i) Lexicon based techniques – These techniques refer to an opinion word list, which contains words that have been identified to express either positive or negative opinions. Given a text, such words are counted and their respective weights are taken into consideration to find the polarity of statements. These word lists are created or compiled using certain seed words, which are expanded to include more words till a standard word list is ready, which can then be referred by texts to perform sentiment analysis. The compilation can be done either done manually, or by automated methods, which are generally dictionary based or corpus based methods [50]. Figure 3 shows a general representation.



Fig. 3. Lexicon – based method

(ii). Machine Learning techniques – Such techniques use a labeled training dataset to train a classifier. Most machine learning methods treat sentiment analysis as a supervised learning problem, though recent methods have explored semi-supervised approaches as well. Unsupervised approaches are difficult to implement, because they require a huge amount of training data to be for accuracy. Furthermore, unsupervised methods may not always match up with human conclusions regarding the given text. Considering sentiment analysis as a classification problem, supervised techniques are more suitable. But availability of plenty of unlabelled data makes it worthwhile to peruse unsupervised methods as well. Figure 4 shows a general representation.

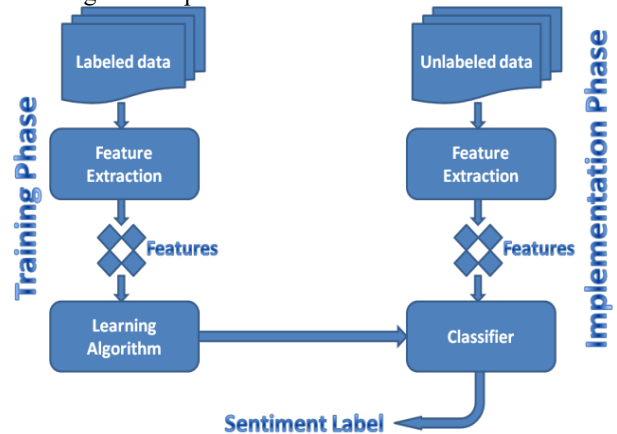


Fig. 4. Machine Learning method

3.2. Performance Measurement Metrics

The sentiment analysis techniques are evaluated using four common metrics, namely Precision, Recall, Accuracy and F-measure. They are defined as follows [13].

$$Precision = \frac{RP}{RP+WP} \quad (1)$$

$$Recall = \frac{RP}{RP+WN} \quad (2)$$

$$Accuracy = \frac{RP+RN}{RP+RN+WP+WN} \quad (3)$$

$$F - Measure = \frac{(2*Precision*Recall)}{(Precision+Recall)} \quad (4)$$

Here, RP = Right Positives, WP = Wrong Positives, RN = Right Negatives and WN = Wrong Negatives.

A confusion matrix represents the respective meanings of the terms, as shown in Table 1.

Table 1. Confusion Matrix for Classification Score

	Calculated Positives	Calculated Negatives
Actual Positives	RP	WN
Actual Negatives	WP	RN

Accuracy signifies the fraction of correct classifications out of the total number of data items provided.

Precision represents the fraction of positive patterns correctly classified out of the positive data items provided.

Recall is the fraction of positive patterns correctly classified out of the total data items provided.

F-measure gives the harmonic mean of the precision and recall values [51].

4. Multi-Lingual/ Cross-Lingual Sentiment Analysis

The bulk of research done on Sentiment Analysis till date has been on English language. In contrast, work on other languages is limited. Of these, a lot of work has been done on Chinese, Japanese and German texts. The approaches to sentiment analysis in other languages have been either multi-lingual or cross-lingual. In multi-lingual methods, specific tools and resources have been developed for the language under research. In cross-lingual methods, existing resources and tools in English are used to perform sentiment analysis in the required language using translation. Cross-lingual methods are more popular because of the plethora of resources and existing research easily available in English, and also because of the newer and faster techniques being developed for automated machine translation [10].

Certain issues affect research in multi-lingual sentiment analysis. The major task is to create reliable and extensive corpus of labeled resources and tools that reflect the intricacies native to the language. This is a time-consuming as well as a labour-intensive task. Because of the extensive resources already built in English, it is an easier approach to first translate to English and then perform sentiment analysis. But this process has its own drawbacks, like not being able to reflect the intricacies of the source language which might have a different sentiment than is obvious from the translation. Some words might also have multiple meanings which get lost in translation. On the other hand, if sentiment analysis in the source language is attempted, there is a requirement to have an exhaustive corpus of labeled resources in that language, which is a time-consuming task. Besides, another issue with multi-lingual sentiment analysis is not handling the domain specific sentiments and biases, which can lead to a drop in the accuracy of the analysis task [11, 12].

4.1. Traditional Approaches to Multi-Lingual and Cross-Lingual Sentiment Analysis

Notable early work on multi-lingual and cross-lingual sentiment analysis includes the NTCIR6 pilot project conducted in Japan in 2007, which has created annotated corpus in Chinese and Japanese languages for research purposes [14]. This corpus was used in the NTCIR7 conducted in 2008, which conducted detailed subtasks like

opinion holder extraction, polarity determination, sentence and clause-level annotation, etc. This work also used a new target language, i.e. Simplified Chinese [15]. In the same year, Boiy and Moens [16] worked on machine learning as a supervised task to reduce the number of annotated examples that can be used for accurate sentiment prediction in French and Dutch. It can be considered as one of the earliest attempts at using machine learning in non-English languages.

During the same time, Wan [17] has presented an experiment that translates Chinese product reviews to English and makes use of the vast English resources for sentiment analysis, which performed better than directly working on the Chinese reviews, which worked on limited data. Machine translation was used, and a combined approach was also tested which proved to be highly effective. Denecke [18] worked on product reviews in German that used the now-popular SentiWordNet corpus to work on the translated reviews, which proved to be a feasible method. This paved the way for further research in the field of multi-lingual and cross-lingual sentiment analysis.

In 2009, Wan [19] carried on his work on Chinese text, wherein he followed a co-training approach, using labeled English data translated to Chinese and unlabeled Chinese data translated to English in combination. This approach worked better than standard classifiers. Brooke et al. [20] advocate the building of resources native to the languages as a better long-term solution, rather than going for automated approaches.

4.2. Multi-lingual and Cross-lingual sentiment analysis using Machine Learning Techniques

Although various methods have been adopted for this task, machine learning techniques remain the most popular and widely-used method. The following section discusses the various techniques and studies that have been conducted for multi-lingual and cross-lingual sentiment analysis using machine learning in recent times. Use of machine learning for multi-lingual sentiment analysis picked up popularity from 2010, due the ease of automation and training of the machine learning systems, and the reduced requirement of manual labour. Some of the techniques employed have been detailed as follows.

Joshi et al. [21] worked on developing an annotated corpus for movie reviews in Hindi language. They have three approaches; namely training a classifier on the Hindi corpus for classification, translation of Hindi to English for sentiment analysis, and development of a lexical resource for score-based classification of Hindi documents. They also describe a fall-back approach by a combination of all three approaches, concluding that the first method performs the best. This method doesn't consider Word Sense Disambiguation and suffers from wrongly translated named entities. Wei and Pal [22] focus on reducing the noise produced by machine translation of Chinese reviews into English using Structural Correspondence Learning. A classifier is used thereafter, which produces better accuracy than the previous co-training approach. This method can improve if it incorporates translator confidence, which it doesn't take into account.

Boyd-Graber and Resnik [23] work on German movie reviews, both on German-English and German-Chinese corpus using supervised Dirichlet allocation, that uses theme-based identification of topics mapped to a ratings variable. It is a novel approach that works on structural connections across languages and also performs Word Sense Disambiguation. The mechanism can be improved by considering local syntax and including more words in its

bridges. Das and Bandyopadhyay [26] have conducted sentiment analysis on Bengali text using both Conditional Random Field and Support Vector Machines are compared their performances using an annotated blog corpus. This method has only been applied on small corporuses, and working on larger corporuses could greatly improve performance.

A lot of research has come forth in this area in 2011. He [24] presents a weakly-supervised technique that uses a latent sentiment model and that considers sentiment labels as topics. The experiment is done on Chinese reviews and the accuracy is found superior to the supervised classification methods. This method has an issue that it depends on the quality of machine translation and is affected by the gap between the source and target language. Pan et al. [25] have also worked on Chinese data by proposing a bi-view matrix tri-factorization model, which then uses the Maximum Entropy classifier on three different domains and produces accurate results. This mechanism uses manual estimation of parameters, and can perform better with a method to validate and set parameters. Xu et al. [27] experiment with two transfer learning algorithms which enhance the training data set and produce an improved accuracy in a SVM classifier, on the NTCIR-7 dataset of Chinese words. Their technique suffers from over-discard of training examples, which can be reduced by a better weighting scheme to improve performance. Tromp, in his thesis [28], compares NB classifier, SVM classifier and a proposed Prior Polarity classifier on social media text, using a Rule-Based Estimation Model Algorithm. He proves that automated sentiment analysis cannot replace traditional surveying, and the proposed mechanism doesn't perform as good on fine-grained sentiment analysis due to lack of extensive training data. Gînscă et al. [29] propose a combination of rule-based classifier, statistical and machine learning methods to implement a sentiment analysis system known as Sentimatrix. This uses a posterior NB classifier and works on Romanian data. Their method can perform better with a more comprehensive resource list, and include modifiers.

In 2012, Wan [30] conducted a comparative study of the different multi-lingual and cross-lingual methods on Chinese sentiment analysis, by taking combinations among them, and shows that an ensemble of the methods give a higher accuracy. He has also provided a different measure of the scoring technique. This experiment proves the inadequacy of individual methods and the strength of combinatorial models. Meng et al. [31] uses a Cross-Lingual Mixture Model which gives better coverage of vocabulary, using a Naive Bayes classifier, both in case of availability or unavailability of labelled data in the target language, i.e. Chinese. The method can do better by expansion of vocabulary. Balahur and Turchi [32, 33] performed experiments on French, German and Spanish data using Machine Translation on three different engines, and employed an SVM classifier, to prove that machine translation gives comparable results to native corpus in the target language. The results are restricted by translation quality and the introduction of noise by addition of translated data together.

Demirtas and Pechenizkiy [34] explore the effectiveness of Machine Translation of Turkish data and claim that increase in size of training data doesn't necessarily improve on accuracy of the classification, using a Naive Bayes classifier, and increase in accuracy can rather be brought about by semi-supervised learning. It doesn't show significant improvement in accuracy with subsequent iterations, which can be worked upon. Volkova et al. [35] explored gender differences on Twitter data, both in Spanish and Russian. The

results are obtained using various classifiers and compared, showing noticeable improvement by including gender demographics. Better results can be achieved by including age, user preferences, etc. These are the major works in 2013.

2014 has been vital for multi-lingual and cross-lingual sentiment analysis. Lin et al. [36] developed a Cross-Lingual Joint Aspect Sentiment model that simultaneously checks aspect-based opinion expression in both languages. They have worked on hotel reviews in a variety of languages, used SVM classifiers, and found it to be more effective with higher accuracy. The method can perform better with availability of a good amount of data. Hajmohammadi et al. [37] used a co-training approach using active learning in combination with semi-supervised learning to provide and enriched training data by picking best samples from unlabelled data, on book review datasets of three languages, to provide an improved performance. The method is limited by the inability of translated data to deal with idiosyncrasy of vocabularies and uses an increased number of features. Again, Lin et al. [38] also proposed a framework that extracts key sentences from text and reduces dependence on external sources, by using a self-supervised learning classifier, on French and German text. This method doesn't show significant improvement over baseline, but is more cost efficient by not using manually annotated corpora. Hogenboom et al. [39] established the significance of relationships between semantics of different languages and map the sentiment scores, achieving higher performance on Dutch input data. This mechanism can improve performance with a more optimized seed set.

A multilingual sentiment elicitation system was developed by Xie et al. [40] which worked on Facebook data, including emoticons, and combined individual techniques that used existing knowledge bases for labeling non-English text for analyzing sentiments with higher accuracy. The technique handles each language independently and can perform better by using a collaborative approach. Xiao and Guo [41] improvised a semi-supervised learning mechanism that induced inter-lingual features and implemented it on four different languages and eighteen classification tasks to achieve significant improvements over other methods. Klinger and Cimiano [42] created a corpus for German and English product reviews, known as USAGE, and performed classification of sentiment at the aspect level using this corpus. The performance is lower for German due to limited informative features available. This can be improved upon. Solakidis et al. [43] also applied a semi-supervised approach that automatically collects training data, using emoticons and emotionally-rich keywords, and applied it on Greek user-generated documents. The mechanism can improve considerably by following a combinatorial approach of keywords and emoticons and using more variety of feature vectors.

Hajmohammadi et al. [44], in 2015, combined active learning along with semi-supervised approach of self-learning, again on book review datasets, while incorporating documents from the target language, and experimented on three different languages, enhancing the performance. The variation in results was due to varying translation qualities. Chen et al. [45] proposed a knowledge validation model for Chinese data that identified credible knowledge, using a semi-supervised technique CredBoost. Experimenting with more knowledge representations can lead to better sentiment information representation.

In 2016, Mozetič et al. [46] conducted experiments of Twitter in various languages and compared the performance of the top classifier models to infer that the size and quality of

datasets impacts the performance more than the selection of the model, and also the important role of the inter-annotator agreement for large training sets. Vilares et al. [47] worked on English and Spanish tweets and compared three varieties of models, out of which the best performance was obtained by a multilingual model trained on a dataset which is multilingual too, and is created by combining monolingual available resources. It also considered the appearance of code-switching texts. The model can work well with better annotations and presence of more Spanish terms. Code switching texts pose additional challenges in analyzing subjectivity.

In 2017, Bhargava and Sharma [59] relied on text summarization to extract important information which is utilized for sentiment analysis, and achieved comparable results. This method can benefit from transliteration and by considering local languages. Becker et al. [60] perform emotion classification using stacked monolingual classifiers and evaluate the effects of machine translation to produce cross-lingual data. The work is extremely comprehensive but can be even more generalized.

As deep learning methods are currently on the rise, newer and more efficient deep learning techniques have also been applied to sentiment analysis. A spurt in this field of work is seen in 2017. Lu and Mori [52] developed a deep-learning based parameter sharing model on a Convolutional Neural Network (CNN) that takes transformed word embeddings and showcases remarkable adaptability over multiple languages. The performances showed non-uniform changes for the same vector space transformation, which can be addressed. Becker et al. [53] also used a similar approach, but used a character based embedding, thereby eliminating translation, and evaluated four combinations of neural models to achieve standard results. Nguyen and Nguyen [54] used a convolutional n-gram BiLSTM method of word embedding that works on YouTube data and provides accurate sentiment analysis on multi-lingual data. The method gave a lower performance on the negative class, due to additional challenges posed by it. Medrouk and Pappa [55] also proposed a convolutional network based model which works on n-gram level information and gives accurate polarity without much prior knowledge. The model performance is affected by quantity of training data, and the performance can be improved by better translation mechanisms. Wehrmann et al. [56] designed a Convolutional Network that uses character-level embeddings and is language-agnostic and free of translation, as well as taking up much lesser memory. The technique is a highly efficient one, and can perform even better with larger corpora as its base. Chen et al. [57] experimented on a different approach that explores document-level semantic connection and works on parallel sentiments, using a Bi-View CNN that gives accurate and stable results. Smadi et al. [58] performed extraction of aspects followed by sentiment analysis on Arabic hotel reviews using two LSTM models and achieved significant improvements over baseline methods. Training on a variety of sentiment lexicons can help the model perform better. Deriu et al. [61] worked on weakly supervised data and experimented with variations of CNN that worked on word embeddings and performed well on sentiment analysis. The mechanism performed better by a single-language approach than a multi-language one, which gives scope for improvement, albeit having other advantages like scalability and adaptability. Peng et al. [62] have followed a different approach towards aspect-oriented sentiment analysis on Chinese data at three granularity levels using a late fusion

technique based on CNN and LSTM. This method could benefit with improvement in the radical level word embeddings which have a scope for better performance.

In 2018, García-Pablos et al. [63] implemented an unsupervised mechanism of topic modeling that enabled aspect-level sentiment analysis for any given language. They combined LDA and word embeddings for bootstrapping, and the system automatically separated opinion words and aspect terms. The method can improve by negation, multi-word and stopword handling. Chen et al. [64] used emojis to bridge between languages and provide a text representation framework, which serve as a distant supervision technique on a bi-directional LSTM model. They perform efficient cross-lingual sentiment analysis by translation and work on an attention model for classifier training. Can et al. [65] utilized RNN to implement cross-lingual analysis by phased training on domain specific models and then applying on translated data. The model is useful for limited data availability and performs well as a generic model. The performance can improve with better quality of translation. Akhtar et al. [66] leveraged bilingual word embeddings and used LSTM for aspect level sentiment analysis successfully on resource-poor Hindi language data. Wang et al. [67] implemented an adversarial cross-lingual learning framework on a CNN architecture that made up for data insufficiency by extracting language-specific and language-independent features that worked effectively on Chinese data. The performance of the model can improve by better translation accuracy. Konate and Ruiying [68] developed six deep learning architectures based on CNN and LSTM to work on Bambara-French social media text by using dictionaries of character and word indexes, which outperformed baseline models. The method can improve performance by working on larger and more exhaustive corpora.

Zhang et al. [69] proposed an auxiliary and attention network based on LSTM, which trained on both resource-rich and resource-poor data. It captured informative words and performed emotion classification by layer sharing. The technique outperformed baseline methods. It can improve by incorporating multiple languages with limited resources. Dong and Melo [70] proposed a cross-lingual propagation algorithm that generated adaptable multi-dimensional sentiment embedding vectors, which was then applied on a dual-channel CNN to perform sentiment analysis on over fifty languages, and five domains. Wehrmann et al. [71] used an architecture based on ConvNet which worked on character based embeddings and performed sentiment analysis as well as language detection, working on less number of parameters. This provided robustness in case of noisy data, saved memory and worked fast. The method does not consider distinct alphabets of specific languages, which could enhance its performance. Rasooli et al. [72] used projection and direct transfer approach to build cross-lingual transfer systems without the need of machine translation, which gave results comparable to supervised approach. It works well in the absence of a dictionary.

Table 2 gives a summary of the various methods discussed.

The pie-chart in Figure 5 gives the percentage of papers that have worked upon various languages to perform multi-lingual and cross-lingual sentiment analysis.

4.3. Challenges of multi-lingual and cross-lingual sentiment analysis

All Sentiment Analysis tasks, including multi-lingual/ cross-lingual sentiment analysis, face a number of issues. Some of them are discussed as follows [4, 6]

- (i). All the five factors, viz. Opinion, Features, Object, Opinion Holder and Time of Expression, may not be available or clearly expressed in the text.
- (ii). All sentences in a text do not necessarily express opinion. Subjectivity classification is needed to identify subjective and objective sentences. It is notable here that some objective sentences may also express opinion, like “The camera stopped functioning within a month.” This clearly states negativity. Such statements express *Implicit Opinion*, as opposed to *Explicit Opinion* stated obviously by subjective sentences. Again, not every subjective sentence expresses opinion; some only express feelings of the opinion holder, like “I supposed the laptop to include accessories at this price.”

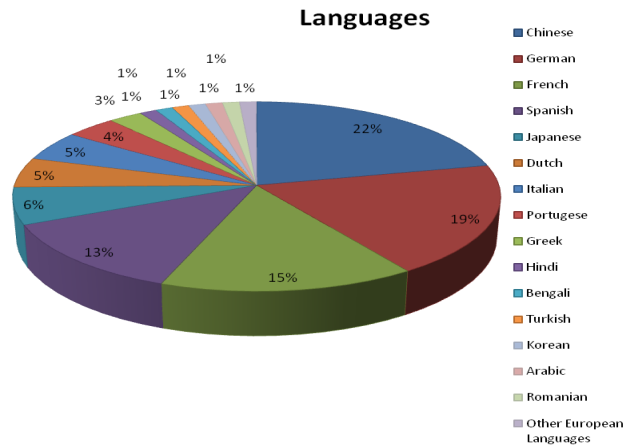


Fig. 5. Languages researched upon

Table 2. Summary of work done on multi-lingual and cross-lingual sentiment analysis

Year	Ref. no.	Method	Target Language(s)	Dataset/ Domain	Major Contributions	Tools Used	Accuracy
2010	[21]	Fall-back strategy; combination of 3 approaches	Hindi	Movie reviews	Three approaches designed - (1) Development of Hindi classifier, (2) Machine Translation of Hindi to English, and (3) development of lexical resource - Hindi SentiWordNet. Combination of the three methods in that order gives best results.	(1) LibSVM, (2) Google Translate, (3) Scoring	(1) 78.14, (2) 65.96, (2) 60.31
	[22]	Structural Correspondence Learning (SCL)	Chinese	Product reviews	Selects reliable parts from translations, SCL finds shared representations between languages	SVM, Google Translate	81.3 - 85.4
	[23]	Multilingual supervised latent Dirichlet allocation (MLSLDA)	German	Movie reviews	Clustering of concepts into coherent topics	GermaNet	MSE : 1.17
	[26]	Fine-grained Sentence Level tagging	Bengali	Blog corpus	Identifying text spans of emotional statements	CRF, SVM	(CRF) 51.00 -68.23, (SVM) 53.35 - 80.55
2011	[24]	Latent Sentiment Model (LSM)	Chinese	Product reviews	Considers sentiment labels as topics and provides preferences to expectations of sentiment labels	Google Translate, Chinese Word Segmenter	81.41
	[25]	Bi-view non-negative matrix tri-factorization model (BNMTF)	Chinese	Book, movie, music reviews	Combines information from two views, includes lexical and label knowledge as well as information from test documents.	Google Translate, Chinese Word Segmenter, Maximum Entropy Classifier	69.3-84.9
	[27]	Transfer AdaBoost and Transfer Self-Training Algorithm	Chinese	Product reviews	TraDaBoost performs automatic adjustment of weights, TrStr trains by enriching with high-quality translated data	SVM, Google Translate	73.04-80.22
	[28]	Four-step approach; Rule-Based Estimator Model Algorithm	German, Spanish, French, Italian, Dutch	News items	Multilingual heuristic approach with high accuracy	SentiWordNet, AdaBoost, Naive Bayes	76.3 - 93.2
	[29]	Sentimatrix - Sentiment analysis service	Romanian	Product reviews	Uses service-oriented architecture for flexible customization, combines rule-based, statistical and machine learning methods	Apache Tika corpus, posterior NB	Precision : 59.8 - 91.2
2012	[30]	Ensemble of various two-dimensional schemes	Chinese	Product reviews	Performs comparative analysis and demonstrates advantages of ensemble approaches, using arbiter-based scoring mechanism	Google Translate, DictTrans	86.1
	[31]	Generative Cross-Lingual Mixture Model (CLMM)	Chinese	MPQA corpus	Learns unknown sentiment words from unlabeled data, utilizes parallel unlabeled data	SVM, Berkeley Aligner	82.7-83.02

	[32]	Machine Translation using three methods on three languages	French, German, Spanish	MPQA corpus	Proves reliability of machine translation by performing comparative analysis	Google Translate, Bing Translator, Moses MT toolkit, SVM with meta-classifier	F-measure : 0.23 - 0.612
	[33]	Machine Translation using three methods on three languages, noise removal using meta-classifier	French, German, Spanish	MPQA corpus	Uses unigrams as features and bagging as meta-classifier to limit noise and sparseness in machine translated data and improve the efficiency	Google Translate, Bing Translator, Moses MT toolkit, SVM with meta-classifier	Recall : 0.703, Precision : 0.752
2013	[34]	Training set expansion and co-training	Turkish	Movie reviews, Product reviews	Studies effect of machine translation on performance, describes which parameters cause improvement	SVM, NB, Maximum Entropy classifier	66-74
	[35]	Bootstrapping lexicons and including hashtags and emoticons	Spanish, Russian	Twitter data	Examines role of gender and usage of emoticons to improve analysis performance	Rule-based classifier	Recall : 0.65-0.73, Precision : 0.68-0.72
2014	[36]	Cross-Lingual Joint Aspect/Sentiment model (CLJAS)	Chinese, German, Dutch, French, Italian, Spanish	Hotel Reviews, Product Reviews	Bilingual aspect-oriented method that uses knowledge from source language by following a topic model framework	Google Translate, HowNet, Universal Sentiment Lexicon (USL), SVM	68.6 - 88.3
	[37]	Active learning and semi-supervised co-training bi-view model	French, German, Japanese	Book reviews	Uses co-training and co-testing to import data from target language, applies combination of active learning and semi-supervised learning, reduces human labelling efforts	Google Translate, Chinese Word Segmenter, MeCab segmenter, SVM	74.24-82.17
	[38]	Self supervised Learning with Key Sentences extraction	French, German	Book, movie, music reviews	Extracts key sentences reflecting sentiment and uses self supervised learning to train without manually labelled corpus	Google Translate	74.3-77.9
	[39]	Semantic lexicon-based sentiment analysis	Dutch	Product reviews	Maps sentiment from source to destination language, based on relations between sentiment lexicons which are language specific, propagates semantics using seed set and related words	SharpNLP POS tagger, Google Translate, Cornetto	62.2
	[40]	Multilingual Sentiment Identification System (MuSES)	Chinese, Korean, German	Product reviews, Facebook and Twitter data	Considers social media semantics, numerically scores phrasal patterns and incorporates emoticons as well as domain knowledge	LingPipe, SVM	76-79
	[41]	Semi-supervised matrix completion	French, German, Japanese	Product reviews	Connects feature spaces of both languages, uses semi-supervised approach on interlingual features	SVM	73.76-83.05
	[42]	USAGE corpus and portability of sentiments	German	Product reviews	Presents a corpus named USAGE and performs aspect-oriented fine-grained sentiment analysis	Classifier	F-measure : 0.32 - 0.76
	[43]	Automatic collection of training data	Greek	Student posts	Includes emoticons and emotionally intense keywords using a semi-supervised approach that excludes human effort	Logistic Regression Classifier	71.2-93.2
2015	[44]	Uncertainty-based active learning and semi-supervised self-training	French, Chinese, Japanese	Book reviews	Combines active learning and self training to reduce human effort, along with automatic labelling, and avoids outlier selection	Google Translate, SVM	70.04-78.63
	[45]	Knowledge validation model (CredBoost)	Chinese	Book, movie, music reviews	Performs knowledge validation in transfer learning and reduces noisy data, uses semi-supervised learning	Google Translate, SVM	80.93-85.18
2016	[46]	Comparative analysis of models under human annotator agreements	Albanian, Bulgarian, German, Hungarian, Polish, Portuguese, Russian, Ser/Cro/Bos, Slovak,	Twitter data, Facebook data	Evaluates datasets using different classifiers by applying different annotator agreements, and determines effects of dataset size on the model performances	NB, Cascading SVM, TwoPlaneSVM, ThreePlaneSVM, TwoPlaneSVMbin, NeutralZoneSVM	46.0-76.0

	[47]	Three methods; Multilingual approach, dual monolingual approach and monolingual language detecting pipeline	Slovenian, Spanish, Swedish Spanish	Twitter data	Designs a code-switching corpus, and compares three approaches, proving the robustness of the multilingual approach, finds no approach performs exceptionally well for code-switching corpus	en-es Classifier, langid.py	52.0-68.7
2017	[52]	Deep learning model with parameter sharing and word embedding	Chinese, Japanese	Twitter data	Unifies parameter sharing and heterogenous word embedding methods in a deep learning model for a multilingual environment	CNN, FastText, Mecab, NLPiR, TweetTokenizer	57.3
	[53]	Deep learning with optimized convolution and character embedding	German, Portugese, Spanish	Twitter data	Uses character based embedding in deep learning models and eliminates the need of machine translation.	CNN, LSTM	66.0 – 69.7
	[54]	Convolutional n-gram Bi-LSTM word embedding	Italian	YouTube comments	Enhances word embedding by multiple convolutions, encodes long distance contextual dependencies.	LSTM	55.03 – 65.6
	[55]	Convolutional nets using n-gram	French, Greek	Restaurant reviews	Uses n-gram level information and works in a language independent manner, excludes code-switching and language translation	CNN	Precision : 0.84 – 0.93
	[56]	Deep CNN with character-level embedding	German, Portugese, Spanish	Twitter data	Language agnostic and translation free analysis depending on fewer parameters, reduces memory usage	CNN	69.7 – 77.0
	[57]	Bi-View CNN (BiVCNN)	Chinese	Book, movie, music reviews	Captures document-level cross-lingual relations in a parallel sentiment space, works on shared polarity between parallel texts	CNN, ICTCLAS, Word2Vec, Google Translate	80.16
	[58]	(a) Bi-LSTM Random Field classifier (b) aspect- based LSTM	Arabic	Hotel reviews	Employs (a) to extract aspect opinion target expression based on n-gram, finds polarity of extracted aspects using (b) using word and aspect embeddings	LSTM, Word2Vec, FastText, AdaGrad	(a) F1 score : 69.98 (b) Accuracy : 82.7
	[61]	Multi-layer CNN for weakly supervised sentiment classification	French, German, Italian	Twitter data	Uses three variants of multilayer CNN on sequences of word embeddings of weakly supervised data and doesn't require translation	CNN	F-measure: 0.63- 0.67
	[62]	Aspect Target Sequence Model (ATSM)	Chinese	Product reviews	Performs multi-grained aspect level sentiment analysis, learns intra-sentence context using word embeddings	LSTM, CNN	75.59 – 5.95
	[59]	Multilingual Sentiment Analysis via Text Summarization (MSATS)	Over 50 global languages	Product reviews	Performs text summarization to extract meaningful information, which is then utilized for polarity detection	SentiWordNet, kNN, SVM, NB, Bing translator	Precision : 0.86
[60]	Multilingual emotion classification	Portugese, Spanish, French	News items	Evaluated effect of translation and language combination on emotion classification, applies stacking of monolingual classifiers	SVM, NB, Radial Basis Function (RBF), Google Translate	F- measure: 0.91-0.95	
2018	[63]	Almost unsupervised system (W2VLDA)	Spanish, French, Dutch	Restaurant reviews, Product reviews	Topic modeling approach that combines LDA with word embedding, separates aspect words and emotion words automatically	Word2Vec, Maximum Entropy Classifier	0.805 (Product), 0.730 (Restaurant)
	[64]	Emoji Powered Representation Learning (Ermes)	Japanese, French, German	Product reviews, Twitter data	Uses emoji-texts to learn representations and finds features for translated documents	Google Translate, MeCab, Word2Vec, LSTM	69.63 – 80.85
	[65]	RNN-based limited data framework	Spanish, Turkish, Dutch, Russian	Product reviews, book reviews, restaurant reviews	Performs general to specific training on RNN on larger data corpus and works on translated data for cross-lingual sentiment analysis	GoogleTranslate, SAS Deep Learning Toolkit, SentiWordNet, RNN	74.36 – 85.61

[66]	LSTM for Multi-lingual and Cross-lingual Analysis	Hindi	Product reviews	Uses bilingual word embeddings trained on a parallel corpus on an LSTM architecture for aspect term extraction and aspect sentiment classification	IndoWordNet, SentiWordNet, LSTM	60.39 - 76.29
[67]	Adversarial Cross-lingual multi-task learning	Chinese	Twitter data	Handles insufficiency of data by adversarial multi-task learning to provide cross-lingual consistency and personalized user sentiment model	Word2Vec, CNN	Precision: 80.02 – 82.03
[68]	Six deep learning models; four LSTM-based and two CNN-based	Bambara-French	Facebook data	Uses fixed indexes to get character and word embeddings and build a dictionary from the code-mixed corpus	Facebook graph API Explorer, CNN, LSTM	82.5 – 83.23
[69]	Aux-LSTM-Attention framework	Chinese	Twitter data	Develops auxiliary representations, uses attention mechanism to capture informative words. Joint learning approach uses labeled data from different language	Baidu Translate, Word2Vec, LSTM, CNN	0.49355
[70]	Cross-lingual propagation algorithm, Dual-channel convolutional neural network	Over 50 languages	Movie, hotel, food, TV series, restaurant reviews	Generates sentiment representations by cross-lingual induction, performs sentiment analysis by Dual-CNN on multiple languages and domains	VADER, SocialSent, GloVe, CNN, SVM	79.30 – 95.92
[71]	Multi-task Neural Network based on ConvNet	German, Portugese, Spanish	Twitter data	Character based approach for feature representation that performs multilingual sentiment analysis and language detection, uses fewer parameters	CNN	73.55 – 74.43
[72]	Cross-lingual sentiment transfer using annotation projection and direct transfer	Arabic, Bulgarian, German, Spanish, Croatian, Hungarian, Polish, Portugese, Russian, Slovak, Slovene, Swedish, Uyghur, Chinese	Twitter data	Uses projection and direct transfer approaches on multiple source languages, includes partial lexicalization, works on comparable data without dictionary availability	SentiPers, GIZA++, LSTM	F-measure: 25.9 – 42.8 (single source), 40.3 – 49.0 (multi-parallel resources)

- (iii). Creating resources is of extreme importance, wherein annotated words are referred to perform the classification. These resources have to be exhaustive and robust so as to provide higher accuracy to the process.
- (iv). Language-specific issues need to be dealt with, like detection of sarcasm and irony, language shortcuts and text-speak, contextual interpretation, etc. which pose difficulties while opinion analysis.
- (v). Spam detection is also a major task, especially when assessing reviews of product. Fake reviews are put up by spammers to promote or defame a product.
- (vi). Word sense disambiguation, i.e. finding the implied meaning of a word having multiple meanings, causes a shift in the sentiment of the text and can confuse the analyzer.
- (vii). The level of annotation also determines the performance of any sentiment analysis algorithm. The annotation could be human or machine generated.

Besides the regular challenges faced by sentiment analysis methods, multi-lingual and cross-lingual analysis methods also have their own set of issues. Some such issues and challenges are described as follows.

- (i). Lack of proper annotated resources – Unlike English, most languages do not have properly annotated and

exhaustive lexical resources, which restricts the training of machine learning models and hence, the models do not achieve their full potential. Any machine learning model gives good performance if it is trained on ample amount of labelled data, the lack of which limits its capability, and hence, most non-English methods still have a scope of performance improvement. As of date, the most exhaustive non-English resources available are of Chinese.

- (ii). Cross-lingual adaptability – Models trained on a specific language do not function well when applied on another language. This happens due to structural and syntactical differences between the languages. For example, if a model is trained on a set of Chinese lexical resources, the model performs well on Chinese but fails to perform equally well when subjected to German. This is owing to the fact that the grammar structure and sentence construction rules are not similar for Chinese and German.
- (iii). Choosing the best method - This entirely depends on the availability of resources and tools already available for the language, including lexicons, dictionaries, processing tools, ease of translation to a corpora-rich language, etc. Each method has its own pros and cons, and suitability of a method is not the same for all types of data or domains. Selecting an apt method depends on both analysis and experimentation.

(iv). Quality of machine translation – The accuracy of the translation mechanism, and the possibility of a sentence losing its structure or meaning after translation, can lead to errors in a translation-based cross-lingual sentiment analysis mechanism. Translation mechanisms, though performing considerably well, have still not achieved near-perfection levels, and are prone to errors in terms of loss of information, mistranslated ambiguous words, and loss of person, voice and tense.

5. Conclusion

In this work, we have studied the available research in multi-lingual sentiment analysis, identified the major languages that have been addressed or which have their own corpus created, enlisted the techniques being used and their contributions, along with their accuracy rates. We have seen the different mechanisms being applied to solve the sentiment analysis problem. Techniques like machine translation are popular in

this field. Work on developing exhaustive corpus in various languages has gained considerable progress.

The concept of multi-lingual sentiment analysis is a lesser explored domain, and has a lot of scope for future work. Even if we have observed substantial amount of work done in certain languages, still there are a number of unexplored languages, which can be considered for further research. Current methods have only an above average accuracy rate, and higher rates may be achieved with the exploration of better techniques and efficient methods. The aspects of speed, reliability, addressing of homonyms and homographs in both source and target languages need to be addressed as well. Hence, this is an open research field for a lot of further work, and has applications in businesses, scientific domains and user awareness.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License



References

1. Aydoğan, E., & Akcayol, M. A. (2016, August). A comprehensive survey for sentiment analysis tasks using machine learning techniques. In *Innovations in Intelligent Systems and Applications (INISTA), 2016 International Symposium on* (pp. 1-7). IEEE.
2. Das, S., & Das, A. (2016, July). Fusion with sentiment scores for market research. In *Information Fusion (FUSION), 2016 19th International Conference on* (pp. 1003-1010). IEEE.
3. Nanli, Z., Ping, Z., Weiguo, L., & Meng, C. (2012, November). Sentiment analysis: A literature review. In *Management of Technology (ISMOT), 2012 International Symposium on* (pp. 572-576). IEEE.
4. Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems, 89*, 14-46.
5. Pawar, A. B., Jawale, M. A., & Kyatanavar, D. N. (2016). Fundamentals of Sentiment Analysis: Concepts and Methodology. In *Sentiment Analysis and Ontology Engineering*(pp. 25-48). Springer, Cham.
6. Montoyo, A., MartiNez-Barco, P., & Balahur, A. (2012). Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments.
7. Hailong, Z., Wenyan, G., & Bo, J. (2014, September). Machine learning and lexicon based methods for sentiment classification: A survey. In *Web Information System and Application Conference (WISA), 2014 11th* (pp. 262-265). IEEE.
8. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal, 5*(4), 1093-1113.
9. Madhoushi, Z., Hamdan, A. R., & Zainudin, S. (2015, July). Sentiment analysis techniques in recent works. In *Science and Information Conference (SAI)* (pp. 288-291).
10. Banea, C., Mihalcea, R., & Wiebe, J. (2011). Multilingual sentiment and subjectivity analysis. *Multilingual natural language processing, 6*, 1-19.
11. Khodier, M. Sentiment Analysis and Opinion Mining in Multi-Language Digital Communities: a Survey.
12. Demirtas, E., & Pechenizkiy, M. (2013, August). Cross-lingual polarity detection with machine translation. In *Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining* (p. 9). ACM.
13. Hailong, Z., Wenyan, G., & Bo, J. (2014, September). Machine learning and lexicon based methods for sentiment classification: A survey. In *Web Information System and Application Conference (WISA), 2014 11th* (pp. 262-265). IEEE.
14. Evans, D. K., Ku, L. W., Seki, Y., Chen, H. H., & Kando, N. (2007, July). Opinion analysis across languages: An overview of and observations from the NTCIR6 opinion analysis pilot task. In *International Workshop on Fuzzy Logic and Applications* (pp. 456-463). Springer, Berlin, Heidelberg.
15. Seki, Y., Evans, D. K., Ku, L. W., Sun, L., Chen, H. H., Kando, N., & Lin, C. Y. (2008, December). Overview of Multilingual Opinion Analysis Task at NTCIR-7. In *NTCIR*.
16. Boiy, E., & Moens, M. F. (2009). A machine learning approach to sentiment analysis in multilingual Web texts. *Information retrieval, 12*(5), 526-558.
17. Wan, X. (2008, October). Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 553-561). Association for Computational Linguistics.
18. Denecke, K. (2008, April). Using sentiwordnet for multilingual sentiment analysis. In *Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on* (pp. 507-512). IEEE.
19. Wan, X. (2009, August). Co-training for cross-lingual sentiment classification. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-volume 1* (pp. 235-243). Association for Computational Linguistics.
20. Brooke, J., Tofiloski, M., & Taboada, M. (2009). Cross-linguistic sentiment analysis: From English to Spanish. In *Proceedings of the international conference RANLP-2009* (pp. 50-54).
21. Joshi, A., Balamurali, A. R., & Bhattacharyya, P. (2010). A fall-back strategy for sentiment analysis in hindi: a case study. *Proceedings of the 8th ICON*.
22. Wei, B., & Pal, C. (2010, July). Cross lingual adaptation: an experiment on sentiment classifications. In *Proceedings of the ACL 2010 conference short papers* (pp. 258-262). Association for Computational Linguistics.
23. Boyd-Graber, J., & Resnik, P. (2010, October). Holistic sentiment analysis across languages: Multilingual supervised latent Dirichlet allocation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing* (pp. 45-55). Association for Computational Linguistics.
24. He, Y. (2011, April). Latent sentiment model for weakly-supervised cross-lingual sentiment classification. In *European Conference on Information Retrieval* (pp. 214-225). Springer, Berlin, Heidelberg.
25. Pan, J., Xue, G. R., Yu, Y., & Wang, Y. (2011, May). Cross-lingual sentiment classification via bi-view non-negative matrix tri-factorization. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 289-300). Springer, Berlin, Heidelberg.
26. Das, D., & Bandyopadhyay, S. (2010). Labeling emotion in Bengali blog corpus—a fine grained tagging at sentence level. In *Proceedings of the Eighth Workshop on Asian Language Resources* (pp. 47-55).
27. Xu, R., Xu, J., & Wang, X. (2011, June). Instance level transfer learning for cross lingual opinion analysis. In *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis* (pp. 182-188). Association for Computational Linguistics.

28. Tromp, E. (2012). *Multilingual sentiment analysis on social media*. Lap Lambert Academic Publ.
29. Gînscă, A. L., Boroş, E., Iftene, A., TrandabĂţ, D., Toader, M., Corîci, M., ... & Cristea, D. (2011, June). Sentimatrix: multilingual sentiment analysis service. In *Proceedings of the 2nd workshop on computational approaches to subjectivity and sentiment analysis* (pp. 189-195). Association for Computational Linguistics.
30. Wan, X. (2012, December). A comparative study of cross-lingual sentiment classification. In *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01* (pp. 24-31). IEEE Computer Society.
31. Meng, X., Wei, F., Liu, X., Zhou, M., Xu, G., & Wang, H. (2012, July). Cross-lingual mixture model for sentiment classification. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1* (pp. 572-581). Association for Computational Linguistics.
32. Balahur, A., & Turchi, M. (2012, July). Multilingual sentiment analysis using machine translation?. In *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis* (pp. 52-60). Association for Computational Linguistics.
33. Balahur, A., & Turchi, M. (2014). Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis. *Computer Speech & Language*, 28(1), 56-75.
34. Demirtas, E., & Pechenizkiy, M. (2013, August). Cross-lingual polarity detection with machine translation. In *Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining* (p. 9). ACM.
35. Volkova, S., Wilson, T., & Yarowsky, D. (2013). Exploring demographic language variations to improve multilingual sentiment analysis in social media. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 1815-1827).
36. Lin, Z., Jin, X., Xu, X., Wang, W., Cheng, X., & Wang, Y. (2014, November). A cross-lingual joint aspect/sentiment model for sentiment analysis. In *Proceedings of the 23rd ACM international conference on conference on information and knowledge management* (pp. 1089-1098). ACM.
37. Hajmohammadi, M. S., Ibrahim, R., & Selamat, A. (2014). Bi-view semi-supervised active learning for cross-lingual sentiment classification. *Information Processing & Management*, 50(5), 718-732.
38. Lin, Z., Jin, X., Xu, X., Wang, Y., Tan, S., & Cheng, X. (2014, August). Make it possible: multilingual sentiment analysis without much prior knowledge. In *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02* (pp. 79-86). IEEE Computer Society.
39. Hogenboom, A., Heerschop, B., Frasinca, F., Kaymak, U., & de Jong, F. (2014). Multi-lingual support for lexicon-based sentiment analysis guided by semantics. *Decision support systems*, 62, 43-53.
40. Xie, Y., Chen, Z., Zhang, K., Cheng, Y., Honbo, D., Agrawal, A., & Choudhary, A. N. (2014). MuSES: Multilingual Sentiment Elicitation System for Social Media Data. *IEEE Intelligent Systems*, 29(4), 34-42.
41. Xiao, M., & Guo, Y. (2014, July). Semi-Supervised Matrix Completion for Cross-Lingual Text Classification. In *AAAI* (pp. 1607-1614).
42. Klinger, R., & Cimiano, P. (2014). The USAGE review corpus for fine-grained, multi-lingual opinion analysis. In *Proceedings of the Language Resources and Evaluation Conference*.
43. Solakidis, G. S., Vavliakis, K. N., & Mitkas, P. A. (2014, August). Multilingual sentiment analysis using emoticons and keywords. In *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02* (pp. 102-109). IEEE Computer Society.
44. Hajmohammadi, M. S., Ibrahim, R., Selamat, A., & Fujita, H. (2015). Combination of active learning and self-training for cross-lingual sentiment classification with density analysis of unlabelled samples. *Information sciences*, 317, 67-77.
45. Chen, Q., Li, W., Lei, Y., Liu, X., & He, Y. (2015). Learning to adapt credible knowledge in cross-lingual sentiment analysis. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (Vol. 1, pp. 419-429).
46. Mozetič, I., Grčar, M., & Smailović, J. (2016). Multilingual Twitter sentiment classification: The role of human annotators. *PLoS one*, 11(5), e0155036.
47. Vilares, D., Alonso, M. A., & Gómez-Rodríguez, C. (2017). Supervised sentiment analysis in multilingual environments. *Information Processing & Management*, 53(3), 595-607.
48. Number of Social Media Users Worldwide 2010-2021 | Statista. Available - <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>
49. Are Blogs a Reliable Source of Information? Available - <https://blog.neongoldfish.com/uncategorized/are-blogs-a-reliable-source-of-information>
50. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267-307.
51. Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 1.
52. Lu, Y., & Mori, T. (2017). Deep Learning Paradigm with Transformed Monolingual Word Embeddings for Multilingual Sentiment Analysis. *arXiv preprint arXiv:1710.03203*.
53. Becker, W., Wehrmann, J., Cagnini, H. E., & Barros, R. C. (2017). An Efficient Deep Neural Architecture for Multilingual Sentiment Analysis in Twitter. *FLAIRS*.
54. Nguyen, H. T., & Le Nguyen, M. (2018). Multilingual opinion mining on YouTube—A convolutional N-gram BiLSTM word embedding. *Information Processing & Management*, 54(3), 451-462.
55. Medrouk, L., & Pappa, A. (2017, November). Deep Learning Model for Sentiment Analysis in Multi-lingual Corpus. In *International Conference on Neural Information Processing* (pp. 205-212). Springer, Cham.
56. Wehrmann, J., Becker, W., Cagnini, H. E., & Barros, R. C. (2017, May). A character-based convolutional neural network for language-agnostic Twitter sentiment analysis. In *Neural Networks (IJCNN), 2017 International Joint Conference on* (pp. 2384-2391). IEEE.
57. Chen, Q., Li, W., Lei, Y., Liu, X., Luo, C., & He, Y. (2017, April). Cross-Lingual Sentiment Relation Capturing for Cross-Lingual Sentiment Analysis. In *European Conference on Information Retrieval* (pp. 54-67). Springer, Cham.
58. Al-Smadi, M., Talafha, B., Al-Ayyoub, M., & Jararweh, Y. (2018). Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. *International Journal of Machine Learning and Cybernetics*, 1-13.
59. Bhargava, R., & Sharma, Y. (2017, January). MSATS: Multilingual sentiment analysis via text summarization. In *Cloud Computing, Data Science & Engineering-Confluence, 2017 7th International Conference on* (pp. 71-76). IEEE.
60. Becker, K., Moreira, V. P., & dos Santos, A. G. (2017). Multilingual emotion classification using supervised learning: Comparative experiments. *Information Processing & Management*, 53(3), 684-704.
61. Deriu, J., Lucchi, A., De Luca, V., Severyn, A., Müller, S., Cieliebak, M., & Jaggi, M. (2017, April). Leveraging large amounts of weakly supervised data for multi-language sentiment classification. In *Proceedings of the 26th international conference on world wide web* (pp. 1045-1052). International World Wide Web Conferences Steering Committee.
62. Peng, H., Ma, Y., Li, Y., & Cambria, E. (2018). Learning multi-grained aspect target sequence for Chinese sentiment analysis. *Knowledge-Based Systems*, 148, 167-176.
63. García-Pablos, A., Cuadros, M., & Rigau, G. (2018). W2vlda: almost unsupervised system for aspect based sentiment analysis. *Expert Systems with Applications*, 91, 127-137.
64. Chen, Z., Shen, S., Hu, Z., Lu, X., Mei, Q., & Liu, X. (2018). Ermes: Emoji-Powered Representation Learning for Cross-Lingual Sentiment Classification. *arXiv preprint arXiv:1806.02557*.
65. Can, E. F., Ezen-Can, A., & Can, F. (2018). Multilingual Sentiment Analysis: An RNN-Based Framework for Limited Data. *arXiv preprint arXiv:1806.04511*.
66. Akhtar, M. S., Sawant, P., Sen, S., Ekbal, A., & Bhattacharyya, P. (2018). Improving Word Embedding Coverage in Less-Resourced Languages Through Multi-Linguality and Cross-Linguality: A Case Study with Aspect-Based Sentiment Analysis. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 18(2), 15.
67. Wang, W., Feng, S., Gao, W., Wang, D., & Zhang, Y. (2018). Personalized Microblog Sentiment Classification via Adversarial

- Cross-lingual Multi-task Learning. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 338-348).
68. Konate, A., & Du, R. (2018). Sentiment Analysis of Code-Mixed Bambara-French Social Media Text Using Deep Learning Techniques. *Wuhan University Journal of Natural Sciences*, 23(3), 237-243.
69. Zhang, L., Wu, L., Li, S., Wang, Z., & Zhou, G. (2018, August). Cross-Lingual Emotion Classification with Auxiliary and Attention Neural Networks. In *CCF International Conference on Natural Language Processing and Chinese Computing* (pp. 429-441). Springer, Cham.
70. Dong, X., & de Melo, G. (2018). Cross-Lingual Propagation for Deep Sentiment Analysis. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI 2018)*. AAAI Press.
71. Wehrmann, J., Becker, W. E., & Barros, R. C. (2018). A Multi-Task Neural Network for Multilingual Sentiment Classification and Language Detection on Twitter. *Machine translation*, 2(32), 37.
72. Rasooli, M. S., Farra, N., Radeva, A., Yu, T., & McKeown, K. (2018). Cross-lingual sentiment transfer with limited resources. *Machine Translation*, 32(1-2), 143-165.