Challenges and Opportunities in Multilingual Sentiment Analysis: Beyond English

Dr Bharathi. V , Assistant Professor, Department of Computer Science(PG), Kristu Jayanti College,Bengaluru.
Bharathi.v@kristujayanti.com
Dr.R.Lakshmi Devi, Assistant Professor, Department of Computer Applications, Women's Christian College, Chennai
Lakshmidevir@wcc.edu.in
Ephraim Godfrey, Sohan Immanuel , Sophy Jose ,
Sohan Immanuel ,Department of Computer Science(PG), *Kristu Jayanti College,Bengaluru*

23mdts17@kristujayanti.com, 23mdts49@kristujayanti.com, 23mdts51@kristujayanti.com

Abstract

With the advent of the World Wide Web, individuals have made extensive use of blogs, social media, and website comments to voice their opinions about a wide range of topics. It becomes clearer and clearer how complicated and advantageous multilingual sentiment analysis of social media is. Sentiment analysis research is expanding incredibly quickly. Due to the wide range of cultural and linguistic backgrounds on the web, analysis of sentiment in English alone is not adequate. This has a big influence on the study of social media and the use of social listening. Multilingual sentiment analysis assists businesses in overcoming language barriers and capturing priceless insights in real-time by realizing that sentiment is inextricably tied to language and culture. This paper is a review on multi-language sentiment analysis, which was presented by many studies published over the last decade addressing the difficulties and possibilities.

Keywords*: Sentiment Analysis, Multilingual Sentiment Analysis (MSA), Social Media.*

1. Introduction

As social media and other online platforms become increasingly embedded in our daily lives, there is an abundance of information available on the internet. As a result, information abounds on all subjects, including goods, companies, market trends, etc. In line with this, a wealth of freely accessible user opinions is readily accessible on the relevant sites. For anyone looking to make an informed decision, it is simple to obtain information on financial market sentiments, political perspectives, movie reviews, product reviews, and other topics.

This kind of scenario calls for the use of a system that helps support decision-making by sifting through the large amount of data, analyzing it, and, ideally, scoring, quantifying, or categorizing it. Opinion mining, another name for sentiment analysis, is a result of this. The process of categorizing, ranking, or calculating the points of view and opinions expressed in materials is known as sentiment analysis. The main objective is to gainan understanding of people's general opinions and feelings towards a certain subject.

Most sentiment analysis research is done on English data, despite substantial progress in thefield. In contrast, very little research has been done on languages in addition to English.

Non-English data analysis is more significant and involves a number of challenges. Other language mechanisms either prefer to translate to English and use the wealth of English resources, or they rely on their own restricted resources. MSA is still in its early stages of development and has a lot of potential for growth.

When English is the primary language and other languages are contributing languages, MSA techniques already work well for language datasets with lots of resources; however, they perform poorly for languages with fewer available resources. The high performance of thesesentiment analysis techniques depends on the availability of labeled datasets or the ability toadapt approaches that deal with the problems associated with under-resourced languages.

Sentiment analysis has been a popular study topic, and many different strategies have been developed to address its challenges. These techniques can be divided into two major categories:

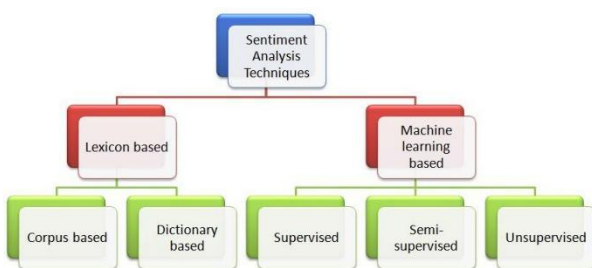

Figure 1. Sentiment Analysis Techniques [10]

Despite the abundance of research on high-resource languages from a multilingual setting, none of them specifically address languages with inadequate resources This paper focuses on presenting a literature review from the viewpoint of languages with limited resources. We have read studies that examined the development of various models for sentiment analysis in numerous languages using machine learning techniques. Additionally, this studyaims to identify existing patterns and constraints in MSA methods for deficient resource languages.

The layout of the paper is as follows. The relevant work on the fundamentals of sentiment analysis is further upon in the second subsection. An overview of the methods and approaches used is provided in the third subsection. Conclusions in the fourth subsection The work is concluded in the fifth subsection with some suggestions for future research.

2. Review of literature

A lot of businesses rely on customer feedback when making decisions. They gather it through the use of questionnaires, customer surveys, opinion polls, etc. Thanks to the internet's ubiquitous availability, a wider range of customers may now be reached by these means of gathering feedback, and they can offer their frank and objective comments. The entire procedure is quicker and simpler. Therefore, it assists companies, suppliers ofservices, online retailers, governments, etc. in gathering and utilizing a variety of viewpointsin order to facilitate decision-making. Ravi K et al, Ravi V et al. (2015).

According to Deriu et al. Sentiment analysis techniques developed for monolingual texts were unable to be used for new or various language sources. Consequently, developing sentiment analysis models that support many languages requires a concentrated effort. For instance, using a machine translation (MT) programme, some researchers offered a multilingual sentiment analysis method, and then they used machine learning (ML) techniques.

The application of MT-based systems for knowledge transfer from resource-rich languagesto under-resourced languages was investigated in earlier MSA studies. These methods convert text to English from a language with limited resources, or conversely, and after thatuse ML-based methods to execute classification of sentiment. Furthermore, this approachtypically has drawbacks like meaning loss and inadequate translations. A. Balahur et al, and M. Turchi et al, 2012.

Wan et al. (2008) implemented machine translation to classify the sentiment of Chinese textsusing sentiment-annotated English text data.

Demonstrating a method that combines a classifier using SVM with features that are independent and dependent on language,Tellez et al,2017.These machine learning methodsadditionally involve a feature extraction step, which we'll eliminate by integrating a deep learning method that handles the aspect of feature learning remotely.

3. Techniques and Methodologies

In this paper we started the literature review by choosing digital libraries and pertinent scholarly databases. ACM Digital Library, IEEE Xplore, PubMed, Google Scholar, and Scopus were among the resources that were selected. These platforms were chosen to guarantee a thorough search across the literature in linguistics, computer science, and naturallanguage processing.

A methodical and organized search approach was developed to find relevant papers for the subject. To guarantee a comprehensive search, keyword combinations such as "multilingual sentiment analysis," "sentiment analysis opportunities," "sentiment analysis challenges," and language-specific terms were employed. Based on predetermined inclusion and exclusion criteria, papers were chosen. Papers discussing potential, difficulties, and multilingual sentiment analysis in the setting of various languages made up the inclusion criteria.

There was a two-step screening procedure used. Titles and abstracts were initially vetted for relevancy. Full-text publications were examined in the second phase to see if they qualifiedfor the review. In order to find common patterns, difficulties, and opportunities in multilingual sentiment analysis, we examined the retrieved data. A thematic synthesis method was used to group the literature and make links between various pieces.

The models employed in the base papers reviewed encompass:

1. Machine Learning: Machine learning techniques in sentiment analysis often include training a classifier with labeled data. Although sentiment analysis is typically approached as a supervised learning task, recent advances have explored semi-supervised methodologies. Using unsupervised techniques can be challenging because large amounts of training data are needed to guarantee accuracy. Furthermore, unsupervised methods may not consistently align with human text interpretations. Supervised techniques are particularly fitting for treating sentiment analysis.

2. CNN And RNN: These are two models designed for text comprehension. They employ specialized techniques for word processing and consist of four main components: embedding, convolutional layers, bidirectional RNN, and fully connected layers. Each model incorporates three branches with unique settings. For example, in the first branch, a specialized filter examines words in groups of three, generating 32 distinctive features, which are then consolidated into a list of 112 numerical values.

The outputs from all three branches are merged before reaching a final decision.
This design is highly effective for comprehending textual content and proves valuable for tasks like text classification and sentiment analysis.(Saad Mboutayeb, Aicha Majda1a and Nikola S. Nikolov,2021).

3. The OpenAI-developed Generative Pre-trained Transformer (GPT) , the LLM underlying the chatbot ChatGPT, has the potential to address the drawbacks of machine learning and dictionary approaches for automated text interpretation. Large text datasets from the internet,such Common Crawl and Wikipedia, are used to train GPT. Which makes it very promisingfor finishing tasks involving text analysis across several languages with no more instruction(sometimes referred to as "zero-shot" learning) .

4. Findings

PAPER	LANGUAGE	TECHNIQUES USED	ACCURACY %

Saad Mboutayeb,Aicha Majda,Nikola S. Nikolov[13]	English, Arabic, Spanish, French, Portuguese	CNN	96.11 86.00 94.34 91.00 94.40
Lu, Y., & Mori, T.[15]	Chinese, Japanese	CNN, FastText, Mecab, NLPIR, TweetTokenizer	57.3

Table 1 Accuracy of Multilingual using ML Techniques [13][15]

1. Besides performing exceptionally well in Portuguese, French, and Spanish, the CNN model excels in English. The Arabic collection of tweets contains a greater range of dialects than the other four languages, which may account for the model's 5–10% worseaccuracy for Arabic. Due to the significant differences between these dialects, Google Translate fails when attempting to translate tweets from non-standard Arabic into English.

2. It was observed that upon space transformation, CNN's accuracy rose by 1.4% while LSTM's accuracy fell by 0.6%.This implies that different types of network architecturesmay not always benefit from the same vector space change. When every factor was considered, the CNN model that was given modified word embeddings performed the best.

Table 2. GPT vs. Top-Performing Machine Learning Models[14]

Language	Construct	GPT-3.5 F1	GPT-4 F1	Top-performing model F1	Model type	Year of model
English	Sentiment	0.685	0.633	0.677	LSTM-CNN	2017
Arabic	Sentiment	0.720	0.707	0.610	Naive Bayes	2017
English	Discrete emotions	0.714	0.779	0.785	BERT	2020
Indonesian	Discrete emotions	0.686	0.740	0.795		2020
English	Offensiveness	0.721	0.746	0.829		2019
Turkish	Offensiveness	0.752	0.709	0.826	XLM-BERT	2020
Swahili	Sentiment	0.560	0.488	0.657	Fine-tuned XLM-R	2023
Hausa	Sentiment	0.586	0.399	0.826		
Amharic	Sentiment	0.238	0.609	0.640		
Yoruba	Sentiment	0.506	0.579	0.800		
Igbo	Sentiment	0.580	0.622	0.830		
Twi	Sentiment	0.408	0.505	0.675		
Kinyarwanda	Sentiment	0.574	0.624	0.726		
Tsonga	Sentiment	0.258	0.302	0.607		

The efficiency of GPT-3.5 and GPT-4 in comparison to the best machine learning models as documented in the articles where we obtained the datasets for testing. Top-performing models are outperformed by GPT-3.5 in identifying sentiment in Arabic and English. GPT-4 performs comparably to discrete emotion models, but not better. For identifying sentiment in African languages, more current refined models outperform GPT-3.5 and GPT-4.

Table 3: MSA: A RNN-Based Structure for Limited Information[19]

Dataset	Majority Baseline	Lexicon-based baseline	RNN (Accuracy)
Spanish restaurant reviews 2,045	72.71	70.98	84.21
Turkish restaurant reviews 932	56.97	61.59	74.36
Dutch restaurant reviews 1, 635	59.63	70.52	81.77
Russian restaurant reviews	79.60	67.81	85.61

RNN consistently outperforms majority and lexicon baselines in sentiment analysis across diverse datasets and languages. The unbalanced training set contributes to more false negatives due to a higher number of positive reviews. Statistical tests (Tukey HSD) and effect sizes (Cohen's d) confirm significant differences, with RNN showing a substantial performance advantage. Figure 3 visually demonstrates RNN's significant superiority over baselines.

5. Conclusion

We examined previous studies on multilingual sentiment analysis, determined which main languages have been studied or for which a corpus has been generated, listed the approachesbeing utilized, their contributions, and their accuracy rates.

Using a recurrent neural network (RNN), the study adopts a general-to-specific model building strategy, training on a diverse, larger corpus before refining on a smaller, single- domain dataset. Results demonstrate that RNN approach outperforms majority and lexicon-based baselines. The RNN models aims to tackles the challenge of creating effective multilingual sentiment analysis models without relying on separate resources for each language.

This paper proposes a multilingual sentiment classification approach using machine translation, deep learning, and NLP techniques. Among five models (Bi-LSTM (Bidirectional Long Short-Term Memory networks), Bi-GRU (Bidirectional Gated Recurrent Units), CNN-LSTM, CNN-GRU.), the CNN model achieves the highest accuracy.The CNN model shows over 90% accuracy for European languages and 86% for Arabic. Challenges arise in translating non-standard Arabic. Future work could explore transformer-based models for sentiment analysis on corpora with diverse Arabic dialects.

Compared to current text analysis techniques like dictionary-based approaches and refined machine learning models, GPT may offer several advantages. It requires no training data, exhibits respectable multilingual accuracy, and is user-friendly with minimal code and easy-to-understand instructions. GPT does not always outperform the best machine learning models, optimized models will probably still be useful. However, GPT performs well enoughto be applied in a variety of scenarios where there is a lack of machine learning classifiers. Moreover, the GPT API is concealed behind a barrier, with more expensive models comingout sooner.

It is discovered that no single model consistently provides accuracy for all languages with fewer resources. Every model that is mentioned in this study has shortcomings or performs less accurately for some languages. Subsequent investigations ought to focus on constructing and refining a model that provides uniform precision for all languages with limited resources.

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]. J. Singh, G. Singh, and R. Singh, A review of sentiment analysis techniques for opinionated web text, CSITrans. ICT, 2016.

[4]. E. Aydogan and M. A. Akcayol, A comprehensive survey for sentiment analysis tasks using machine learning techniques, 2016 Int. Symp. Innov. Intell. Syst. Appl., pp. 17, 2016.

[5]. Singh VK, Piryani R, Uddin A, Waila P, et al. Sentiment analysis of textual reviews; Evaluating machinelearning, unsupervised and SentiWordNet approaches. In: 2013 5th international conference on knowledge andsmart technology (KST). IEEE; 2013, p. 122–27.

[6]. Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-Based Systems.

[7]. A. Balahur and M. Turchi, ``Multilingual sentiment analysis using machine translation?'' in Proc. 3rd Workshop Comput. Approaches Subjectivity Sentiment Anal., (2012).

[8]. Wan X (2008) Using bilingual knowledge and ensemble techniques for unsupervised chinese sentiment analysis. In: Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing.

[9]. Eric S Tellez, Sabino Miranda-Jim´enez, Mario Gra Daniela Moctezuma, Ranyart R Su´arez, and Oscar S Siordia.(2017). A simple approach to multilingual polarity classification inTwitter. Pattern Recognition Letters.

[10]. 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.

[11]. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093-1113.

[12]. Madhoushi, Z., Hamdan, A. R., & Zainudin, S. (2015, July). Sentiment analysis techniques in recent works. In Science and Information Conference (SAI) (pp. 288-291).

[13]. Saad Mboutayeb,Aicha Majda,Nikola S. Nikolov,Multilingual Sentiment Analysis: A Deep Learning Approach,January (2021).

[14]. Rathje, S., Mirea, D., Sucholutsky, I., Marjieh, R., Robertson, C., & Van Bavel, J. J. (2023, May 19). GPTis an effective tool for multilingual psychological text analysis.

[15]. Lu, Y., & Mori, T. (2017). Deep Learning Paradigm with Transformed Monolingual Word Embeddings for Multilingual Sentiment Analysis. arXiv preprint arXiv:1710.03203.

[16]. A. Ghafoor, A. S. Imran, S. M. Daudpota, Z. Kastrati, Abdullah, R. Batra, and M. A. Wani, ``The impactof translating resource-rich datasets to low resource languages through multilingual text processing," IEEE Access, vol. , (2021).

[17]. Sagnika, Santwana; Pattanaik, Anshuman; Mishra, Bhabani Shankar Prasad; and Meher, Saroj K., "A review on multi-lingual sentiment analysis by machine learning methods" (2020). Journal Articles.

[18]. Ghafoor, A., Imran, A S., Daudpota, S M., Kastrati, Z., Soomro, A. et al. (2021) The Impact of TranslatingResource-Rich Datasets to Low-Resource Languages Through Multi-Lingual Text Processing.

[19]. Multilingual Sentiment Analysis: An RNN-Based Framework for Limited Data (Ethem F. Can, AysuEzen-Can, Fazli Can).