DETECTING SENTIMENT EXPRESSED IN LOW RESOURCE LANGUAGES: MATERIAL AND METHODS

Supported by tetfund
TERTIARY EDUCATION TRUST FUND

Ibrahim Salim Suleiman*, Sunday Eric Adewumi, Victoria Ifeoluwa Yemi-Peters
Department of Computer Science
Federal University Lokoja, Kogi State
*Corresponding Author: salimsuleimanibrahim@gmail.com

Received: September 14, 2024 Accepted: December 28, 2024

Abstract: Despite the significant attention gained by sentiment analysis in Natural Language Processing (NLP), its application on low-resource languages is limited because of the limitations in their dataset and underdeveloped tools. This literature review addresses these challenges by reviewing recent academic papers to show previous approaches and algorithms used for sentiment analysis in low-resource languages. This study utilized the systematic literature review method that was proposed by Kitchenham et al. (2009), which involves the formulation of three research questions to show the approaches previously used for analyzing sentiment in low-resource languages, the algorithms previously used for sentiment analysis in low-resource languages, and the performance of previously used algorithms for sentiment analysis in low-resource languages. To answer these research questions, 42 relevant papers were extracted from reputable academic databases: Scopus, IEEE Explore, Springer Link, and Web of Science. These papers were extracted using the search strings based on this study's inclusion and exclusion criteria. The papers were carefully analyzed, and the result shows a significant gap in the previous model's adaptability and performance, indicating a need for more frameworks on content awareness and robust sentiment analysis. This research findings give a good foundation for future research which should aim at developing a method that can enhance high accuracy and real-world application of sentiment analysis towards low resource languages. Transformer-based models, like BERT and RoBERTa, achieved up to 99.94% F1-score. Deep learning models, such as BiLSTM, reached 94.6% accuracy, while hybrid approaches like BUQRNN achieved 99.3% accuracy, showcasing significant performance improvements in low-resource languages. Future work should focus on creating and testing innovative algorithms and datasets to bridge the identified gaps and advance sentiment analysis in underrepresented linguistic domains.

Keywords: Natural Language Processing, Sentiment Analysis, Literature Review, Low-Resource Languages, Algorithms.

Introduction

Natural Language Processing (NLP) is a branch of computer science used to understand human languages through computational models. NLP is divided into two main areas which are the core areas and applications areas. The first part which is the core area focuses mainly on fundamental problems like language modeling, morphological processing, syntactic processing, and semantic processing (Cui, et al., 2023). The second part which is the application areas focuses on tasks such as language translation (text), book summarization, information search, text generation, and classifying and clustering documents. To solve the application issues, one or more fundamental problems need to be solved (Li et al., 2023). NLP is mostly based on machine learning techniques along with probabilistic and statistical calculations, using algorithms like support vector machines (SVM), decision trees (DT), random forest (RF), and k-nearest neighbor (KNN) among others. Recently, NLP has gone through a major transformation with the adoption of neural network-based models such as convolutional neural networks (CNN), long short-term memory networks, recurrent neural networks, and residual connections, among others, as adding to the traditional machine-learning models (Cofino et al., 2024). Also, recent researchers on NLP have applied pre-trained models like Bi-directional Encoder Representations from Transformers (BERT) on sentiment analysis, and these models have proven to be effective in analyzing sentiment.

Sentiment analysis is a subfield of NLP, and it is the process of analyzing large amounts of data to detect sentiments such as positive, negative, or neutral sentiments. Researchers in sentiment analysis found it difficult to carry out sentiment analysis on low-resource language (Bu et al. 2024). Many tools used for sentiment analysis such as lexicons, annotated datasets, and pre-trained models are not available for the low-resource language, even with the fact that there is a growing interest in the application of sentiment analysis on low-resource language, and this area is still under-researched. In light of this, this research aims to provide a comprehensive review of the current state of sentiment analysis in low-resource language, focusing on approaches, algorithms, and their performance to show the challenges, and future directions of this research area.

The rest of this research is organized as follows: Methodology, which shows the stepwise method used to carry out this research; Result, which shows the outcome of the research conducted; discussion, which shows the analysis of insight gained from the research's results; and conclusion, which shows the summary of the key outcomes and implications of the study.

Methodology

We adopted an in-depth and reliable systematic literature review method proposed by Kitchenham et al. (2009). This methodology involves formulating three research questions, obtaining relevant research papers based on this study's inclusion and exclusion criteria, and finding answers to the formulated research question by reviewing and analyzing the gathered research papers.

Inclusion criteria

Journal and Conference papers written in the English language, and published from 2020 to 2024 were included in this research. Additionally, these papers must be relevant to NLP, sentiment analysis, and low-resource languages to qualify for inclusion.

Exclusion criteria

Papers that are not written in English and unrelated to NLP, sentiment analysis, and low-resource languages were excluded. Also, papers whose abstracts do not explicitly state contributions to knowledge are excluded from this study's review papers.

Search string

We used the keywords from the research questions to construct a search phrase to create good search strings. The search strings were used for papers in the databases provided in Table 1. The search strings are "sentiment analysis", "low-resource language", "sentiment classification", "sentiment detection", "approach for sentiment analysis", "sentiment analysis models", and "review on sentiment analysis".

Table 1: Academic database

SN	**Database**	**Publisher**	**Link**
1	Scopus	Elsevier	https://www.scopus.com/home/
2	IEEE Explore	IEEE	https://ieeexplore.ieee.org/Xplore/home/
3	Springer Link	Springer	https://link.springer.com/
4	Web of Science	Clarivate	https://www.sciencedirect.com/

Search process

1,000 papers were initially collected from the databases presented in Table 1, using the search strings. To mitigate selection bias, we implemented a transparent and rigorous inclusion criterion, ensuring that studies were selected from a wide range of low-resource languages, geographical areas, and sentiment analysis approaches. We then carried out the first screening of the documents by removing irrelevant ones based on the abstract papers, and 672 papers were removed, leaving 328 papers. We further performed the second screening after going through the content of the paper, where 42 relevant papers were finally extracted for review. The Search and Screening Results from the databases are displayed in Table 2.

Table 2: Search and screening results from databases

	Scopus	IEEE Explore	Springer Link	Web of Science	Total
First search	147	160	383	310	1000
First screening	76	54	96	102	328
Second screening	9	9	11	13	42

Research questions

Research question 1: What approaches were previously used for analyzing sentiment in low-resource languages?
Research question 2: What algorithms were previously used for sentiment analysis in low-resource languages?
Research question 3: How well have the previously used algorithms for sentiment analysis in low-resource languages performed?

Literature Review

Roy (2024) applied a deep learning and transfer learning method to the sentiment analysis method on a bi-lingual low-resource language. The researchers combine the deep learning approach with the transformer approach, which makes it a hybrid method. The dataset that was used in the research consists of a code-mixed Kannada and Malayalam language, which was collected from a different online source, such as users' posts on social media. The specific algorithm that was used is the artificial neural network(ANN) and the Bidirectional encoder representations from Transformers (BERT). The result of the experiment shows that the transformer-based ensemble model on the Kannada Code-Mixed Dataset achieved an F1-score of 66%, while the ensemble machine learning model on Malayalam Dataset achieved 72% accuracy.

Hashem et al. (2024) used the deep learning approach to perform sentiment analysis in low-resource languages. The dataset contains user-generated opinions and reviews in the Arabic language and was collected from the X platform which was formally known as Twitter, the dataset contains Tweets during the COVID-19 pandemic. The research adopted the Bi-LSTM algorithm for sentiment classification. The result shows that the proposed Bi-LSTM achieved an accuracy of 94.6% on the sentiment classification.

Yu et al. (2024) carried out a sentiment analysis on Bengali text which is a low-resource language in South Asia. The researchers employed a hybridization of deep learning and quantum machine learning transfer learning approaches. The model used includes the Batch Upload Quantum Recurrent Neural Network (BUQRNN) for feature extraction and

Parameter Nonshared Batch Upload Quantum Recurrent Neural Network (PN-BUQRNN) for advanced feature extraction. The result shows that the BUQRNN archives 99.3% accuracy outperforming the PN-BUQRNN model.

Onan & Balbal (2024) conducted research that applied integrated universal transformation and task-specific methods on Turkish text dataset augmentation for sentiment analysis. The researchers made use of 8 different pre-trained language models language model which include BERT-Large-ITPT, XLNet, BERTurk, DistilBERTurk, ULM-Fit, SSL, ELECTRA, and ConvBERTurk. The result shows that the models were able to achieve accuracies that range from 86% – 93%, with the BERT-Large-ITPT achieving the highest accuracy of 93%.

Husain et al. (2024) analyzed sentiment in the Kuwaiti language after addressing the shortage of annotated resources with the use of a transformer-based approach for sentiment analysis. The researchers fine-tuned a pre-trained BERT algorithm and called it ARBERT, which was used for the sentiment analysis. The result from the analysis shows that the proposed ARBERT was able to achieve 89% accuracy.

Saroj et al. (2024) conducted research that used the deep learning approach for sentiment analysis in the Hindi (SAFH) language. The dataset which consists of 10,011 tweets was collected from the X platform. The proposed deep learning model was used with the Word Embedding technique. The proposed model was able to achieve 90.9% accuracy.

Baniata & Kang (2024) carried out research that employed a hybrid approach for sentiment analysis in the Arabic language. The hybrid approach combines Transfer learning with Multi-Task Learning (MTL) and the Mixture of Experts (MoE). The sentiment analysis was carried out on three different datasets which were labeled as HARD, BRAD, and LABR. The dataset contains Arabic text across 3 to 5 polarity tasks with over one million entries. The result of the experiment shows that the proposed algorithms were able to achieve 84.02% accuracy on the HARD dataset, 67.89% accuracy on the BRAD dataset, and 83.91% accuracy on the LABR dataset.

Wang & Xiong (2024) carried out a sentiment analysis on a book review in English Language. The researcher proposed the use of a hybrid approach that combines the pre-trained and the deep learning approach. The hybrid algorithm combines BERT for Word Embedding and BiLSTM to learn contextual information (BERT-BiLSTM). The experimental result shows that the proposed model was able to achieve an improved precision of +4.2%, recall of +3.9%, and F1 score of +3.79%.

Cosme & de-Leon (2024) utilized the transfer learning approach on sentiment analysis in the Filipino-English language with code-switching. The dataset that was used is the first sentiment-annotated corpus for code-switching text, which contains over ten thousand reviews. The model used is a fine-tuned XLM-RoBERTa which was able to outperform all other compared algorithms with an accuracy and F1-score of 84%.

Alqurafi & Alsanoosy (2024) carried out a sentiment analysis on a large real-world dataset that was collected from customer reviews on different platforms like eBay or Amazon. The study used different machine learning and deep learning algorithms which include Logistic Regression (LR), Support Vector Machines (SVM), Decision Trees (DT), and LSTM. The result shows that the LSTM model outperformed all other compared algorithms by achieving 93.3% accuracy. While the DT has the worst performance with 77.3% accuracy, and LR and SVM achieved 91.3% accuracy each.

Mohammed & Prasad (2023) applied a hybrid of rule-based and lexicon-based approaches to analyze sentiment in Hausa Tweets Corpus. The study used the Lexicon-Based Sentiment Analysis Model which includes processes like lexicon creation, data augmentation, data annotation, and fine-tuning. The result of the analysis shows that the proposed model achieved 98% accuracy.

Yusup et al. (2023) utilized a hybrid approach that combines TextCNN, RNN, RCNN, and LSTM to analyze sentiment in the Uyghur language. The dataset was collected using syntactic-semantic knowledge from a conversion of a high-resource language corpus to a low-resource corpus. The proposed hybrid model achieved 82.17% accuracy.

Ghasemi & Momtazi (2023) conducted research that utilized the transfer learning approach h for sentiment analysis of a low-resource language. The study used the XLM-RoBERTa algorithm. The experiment's results show that the proposed model achieved a 0.55% improvement in accuracy over cross-lingual models and a 5.08% improvement over monolingual models.

Vo et al. (2023) applied the deep learning approach to sentiment analysis of the Vietnamese language. Vietnamese corpus was collected from the University of Phan Thiet and was divided into positive, neutral, and negative sentiments. The deep learning algorithms used include LSTM, BERT, DistilBERT, BERT-RCNN, and PhoBERT. The experimental result shows that the PhoBERT algorithm was able to achieve the best F1 score of 89.68%, while LSTM, DistilBERT, BERT, and BERT-RCNN achieved F1-scores of 72.11%, 82.40%, 86.25%, and 86.83 respectively.

Rasool et al. (2023) applied a hybrid approach to the sentiment analysis of Asian low-resource language. The hybrid approach combines the Pelican Optimization Algorithm (POA), Deep Learning, and Aspect-Based Sentiment Analysis (ABSA). The ABiLSTM algorithm was used for sentiment classification. The result from the analysis shows that the proposed model achieved 98.72% accuracy and 98.71% precision, showing superior performance when compared to the existing classification models.

Raychawdhary et al. (2023) applied an approach that combines transfer learning and the ABSA to sentiment analysis on the Hausa and Igbo languages spoken in Nigeria. The dataset that consists of tweets in Hausa and Igbo was collected from the X platform. The dataset used consists of code-mixed tweets and has sentiments labeled as positive, negative, and neutral. The study fine-tuned a pre-trained language model called AfriBERTa, which was used for sentiment classification. The experimental result shows that the proposed model achieved 80.85% F1 scores for the Hausa language and 80.82% F1 scores for the Yoruba language.

Hancharova et al. (2023) presented a hybrid approach for sentiment analysis on low-resource languages in Africa. The hybrid method combines Machine learning, Translation-based, and deep learning approaches. The SemEval-2023 Task 12 dataset was used. The SVM model was used for sentiment classification. The result shows that the proposed methods were able to attain an F1-Score of 68%.

Akrah & Pedersen (2023) adopted the transfer learning approach for sentiment analysis of the Twi language. The Twi language dataset was collected from the X platform. The pre-trained model tailored for the Twi language was called TwiBERT. The study's results show that the proposed model achieved a 64.29% F1 score.

Garcia-Diaz et al. (2023a) used an approach that combines Deep Learning with a linguistic feature-based approach for Sentiment Analysis in the Spanish language. The dataset that was used contained 15,915 financial tweets in Spanish. The RoBERTa algorithm was used for sentiment classification. The result shows that the proposed method achieved 73.15% F1-Score.

Salahudeen et al. (2023) applied the Transfer Learning to sentiment analysis of 14 different African languages. The study used 5 different pre-trained algorithms, and the algorithms were fine-tuned on specific low-resource datasets. The algorithms include AfriBERTa-Large, BERT, mBERT, Afro-xlmr-Large, and Arabic-CamelBERT. The result from the analysis shows the accuracies of the AfriBERTa-Large, BERT, mBERT, Afro-xlmr-Large, and Arabic-CamelBERT to be 60.53%, 53.12%, 59.53%, 62.26%, and 56.35% respectively.

Ramanathan et al. (2023) utilized the transfer learning approach on Hausa sentiment analysis. The study collected the ow-resource African language dataset used from Twitter. The Multilingual RoBERTa which was pre-trained in 100 languages was used for sentiment analysis while the was used for tokenization. The result shows that the proposed model achieved 99.94%F1score on the training data and 80.32% F1 score on the test data.

García-Díaz et al. (2023b) applied a hybrid approach that combines deep learning and the rule bases for sentiment analysis in some specific low-resource languages in Africa. The languages include the Xitsonga, Algerian Arabic, Swahili, and Twi. The algorithms used include the Multilingual Large Language algorithm for feature extraction and Language-Independent Linguistics to improve model performance. The result shows that the proposed method achieved F1 Scores of 54.89%, 68.52%, 60.52%, and 71.14% for Xitsonga, Algerian Arabic, Swahili, and Twi respectively.

Prakash & Vijay (2023) applied a hybrid approach that combined transfer learning and deep learning approaches to analyze sentiment from reviews of Tamil movies. The study utilized the Multi-Stage Deep Learning Architecture (MSDLA) algorithm. The result from the research shows that the proposed method achieved 87.72% accuracy and 86.14% precision.

Abdirahman et al. (2023) utilized both machine learning and the deep learning method for the Somali language dataset. The algorithms used include the Decision Tree Classifier, Random Forest Classifier, Extreme Gradient Boosting (XGB), CNN, and LSTM for classification for sentiment classification. The result of the experiment shows that the LSTM model performs best among the deep learning models with the highest accuracy of 88.58%, while the Decision Tree Classifier performs best among the machine learning algorithms with an accuracy of 87.94%.

Mujahid et al. (2023) utilized a hybrid approach that integrates Transfer Learning with the deep learning approach. The study used the hybrid model of a transformer-based model that combines RoBERTa and BERT for sentiment classification. The sentiment analysis was done on the ChatGPT Tweets Arabic Dataset. The result shows that the proposed model was able to achieve 96.02% accuracy.

Ullah et al. (2022) applied an approach that combines Deep Learning models with NLP techniques. The CNN-LSTM paired with the Word2Vec embeddings model was used for sentiment classification on Urdu text. The result from the sentiment analysis shows that the proposed model was able to achieve its highest accuracy of 93.3%.

Altaf et al. (2022) proposed an approach that combines Deep Learning techniques with NLP-based features to analyze sentient in Urdu text. Two datasets from different domains of Urdu text are used to analyze cross-domain sentiment classification. The deep learning algorithms used include RNN, LSTM, and GRU algorithm. The result shows that the GRU model achieved 76% F1-score and 77% accuracy, the RNN model achieved 72% F1-score and 76% accuracy, and the LSTM model achieved 75% F1-score and 76% accuracy.

Shanmugavadivel et al. (2022) applied three different approaches for sentiment analysis and offensive language detection in code-mixed Tamil-English data. The utilized approaches included machine learning, deep learning, and transfer learning techniques. The algorithms used for machine learning include logistic regression, CNN, Bi-LSTM, BERT, RoBERTa, and Adapter-BERT. The result shows that the Adapter-BERT outperformed all other compared algorithms, achieving 65% accuracy on sentiment analysis.

Das & Singh (2022) presented research that utilized a hybrid approach that integrates Deep Learning techniques with multimodal fusion mechanisms, this method combines textual and visual features to perform sentiment analysis for the low-resource Assamese language. The proposed Multimodal Framework achieved 89.3% Accuracy.

Mamta et al. (2022) applied the transfer learning approach to sentiment analysis, leveraging knowledge from a high-resource language (English) to improve sentiment analysis performance for a low-resource language (Hindi). The dataset was gathered from movie and product reviews. The Cross-lingual Word embedding algorithm was used to map the Hindi and English languages to a shared semantic space

for knowledge transfer. The result shows 60.09% accuracy for the movie review dataset and 72.14% accuracy for the product review dataset.
Khan et al. (2022) applied three different approaches to sentiment analysis in the Urdu language, the approaches include machine learning, deep learning, and rule-based approaches. The algorithms that were used are SVM, NB, Adaboost, MLP, LR, and RF for the machine learning approach, CNN, LSTM, Bi-LSTM, GRU, and Bidirectional GRU (Bi-GRU) were used for the deep learning approach while the transfer learning, the pre-trained mBERT was fine-tuned specifically for Urdu sentiment analysis. After comparing the results of all the algorithms used, the pre-trained model was able to outperform all the other compared algorithms, achieving 81.49% accuracy.
Bensoltane & Zaki (2022) applied the transfer learning approach for sentiment analysis on Arabic news posts. AraBERT which is an Arabic version of the BERT algorithm was used to perform an aspect sentiment analysis. The result from the experiment shows that the proposed model achieved more than 80% accuracy, achieving more than 19% accuracy over the compared algorithms.
Garg, K., & Lobiyal, D. K. (2021) applied a hybrid method of deep learning and fuzzy logic for sentiment analysis in Hindi. The neuro-fuzzy algorithm was used to carry out the analysis. The result from the analysis shows that the proposed model achieved an accuracy of 91.18% on the Hindi dataset, outperforming state-of-the-art algorithms like the NB and SVM.
Tho et al. (2021) utilized a hybrid of deep learning and a Rule-Based Approach to sentiment analysis in Bahasa Indonesia and Javanese languages. The dataset that was used is a code-mixed text in Bahasa Indonesia and Javanese. The algorithm used for the sentiment classification is a pre-trained deep learning model named Sentence-BERT. The result of the experiment shows that the proposed model achieved an average of 83% accuracy and 90% precision.
Kastrati et al. (2021) utilized a deep learning approach for sentiment analysis in the Albanian language. The dataset used contained 10,742 manually classified Facebook comments in Albanian. The BiLSTM with Attention Mechanism was used for sentiment analysis. The result shows that the proposed system achieved an F1-Score of 72.09%, outperforming other compared algorithms.
Bhowmick & Jana (2021) applied the transfer learning approach to sentiment analysis in the Bengali language. The algorithms used are the pre-trained Multilingual BERT which was fine-tuned for sentiment analysis, and the pre-trained XLM-RoBERTa which was fine-tuned for the task. The result shows that the proposed model was able to achieve 95% accuracy.
Khamphakdee & Seresangtakul (2021) applied supervised machine learning algorithms for sentiment polarity classification in the Thai language. The dataset that was used is a review in the Thai language which was collected from Agoda.com and Booking.com. Different machine learning algorithms were used, but the SVM outperformed all the compared algorithms, achieving the highest accuracy of 89.96%.
Divate (2021) applied the deep learning method to polarity-based sentiment analysis of Marathi e-news. The Marathi dataset was collected from different e-news platforms. The deep learning algorithm that was used in the study is the LSTM algorithm. The result of the proposed method shows that the LSTM was able to achieve 72% accuracy in identifying sentiment polarity.
Bharti et al. (2020) proposed a Rule-Based Approach to sarcasm detection in Telugu conversational sentences annotated. The proposed algorithm includes TNWS Algorithm, TIWS Algorithm, and RiFoQ Algorithm. The experimental result shows that the TNWS algorithm achieved 91% accuracy, the TIWS algorithm achieved 88.5% accuracy, and the RiFoQ algorithm achieved 91.5% accuracy.
Xu et al. (2020) utilized the transfer learning approach on aspect-based sentiment analysis. The study used the DomBERT algorithm, which combines domain-specific language understanding, and general-purpose language models. The experimental result shows that the proposed algorithm achieved its best performance of 83.14% accuracy on Aspect Sentiment Classification and 73.45% accuracy on aspect-based sentiment analysis.
Singh et al. (2020) applied the deep learning approach for sentiment analysis in the Nepali language. The dataset was collected from YouTube. The BiLSTM algorithm was used for sentiment classification. The result from the experiment shows that the proposed BiLSTM algorithm was able to achieve 81.60%F1-score for sentiment classification.
Aliramezani et al. (2020) applied an approach that combines Cross-Lingual Transfer Learning and Deep Learning techniques for sentiment analysis Persian language. The RNN algorithm was used to build the model. The dataset used is a Persian Snapp Food dataset. The proposed sentiment analysis algorithm was able to achieve an F1 score of 78.16% on Persian test data.

Result

Answer to research question 1

Research question 1**:** What approaches were previously used for analyzing sentiment in low-resource languages?

Table 3: Approaches for sentiment analysis in low-resource languages

Approach	Authors
Machine Learning -based approach	Alqurafi, & Alsanoosy (2024), Hancharova et al. (2023), Abdirahman et al. (2023), Shanmugavadivel et al. (2022), Khan et al. (2022), Khamphakdee & Seresangtakul (2021).
Deep learning-based approach	Roy (2024), Hashem et al. (2024), Onan & Balbal (2024), Saroj et al. (2024), Alqurafi, & Alsanoosy (2024), Vo et al. (2023), Vo et al. (2023), Rasool et al. (2023), Abdirahman et al. (2023), Ullah et al. (2022), Altaf et al. (2022), Shanmugavadivel et al. (2022), Das & Singh (2022), Khan et al. (2022), Kastrati et al. (2021), Divate (2021), Singh et al. (2020), Aliramezani et al. (2020).
Transformer-based approach	Roy (2024), Onan & Balbal (2024), Husain et al. (2024), Cosme & de-Leon (2024), Mohammed & Prasad (2023), Ghasemi & Momtazi (2023), Vo et al. (2023), Vo et al. (2023), Raychawdhary et al. (2023), Akrah & Pedersen (2023), Garcia-Diaz et al. (2023a), Salahudeen et al. (2023), Ramanathan et al. (2023), Mujahid et al. (2023), Shanmugavadivel et al. (2022), Mamta et al. (2022), Bensoltane & Zaki (2022), Bhowmick & Jana (2021), Xu et al. (2020).
Rue-based approach	Mohammed & Prasad (2023), García-Díaz et al. (2023b), Khan et al. (2022), Bharti et al. (2020).
Hybrid approach	Yu et al. (2024), Baniata & Kang (2024), Wang & Xiong (2024), Mohammed & Prasad (2023), Vo et al. (2023), Prakash & Vijay (2023), Mujahid et al. (2023), Ullah et al. (2022), Garg, K., & Lobiyal, D. K. (2021), Tho et al. (2021).

Based on the review papers we found out that the previously used algorithms for sentiment analysis in low-resource languages are the machine learning approach, deep learning approach, transformer-based approach, rule-based approach, and hybrid approach. As shown in Table 3, we observed that the transformer-based approach has been mostly used by researchers over the past 5 years, and the deep learning, machine learning, and hybrid approaches are growing in popularity, while the rule-based methods are limited but notable.

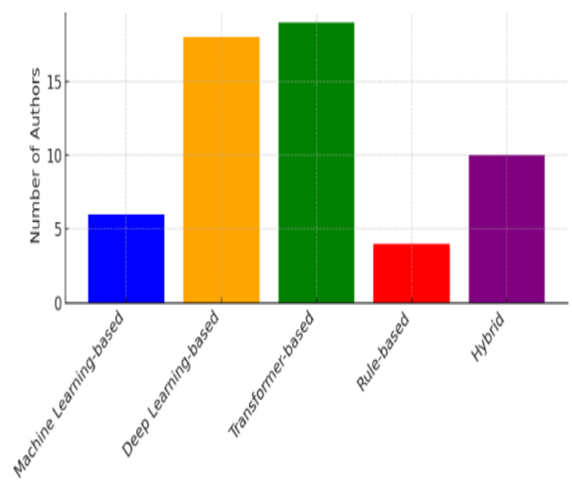

Figure 1: Research contribution by authors for sentiment analysis.

Figure 1 presents a bar chart that illustrates the research contributions by authors for sentiment analysis for low-resource languages. From the chart, we observed that the transformer-based approach and deep learning approach have the highest contributors with 19 and 18 contributions respectively. The hybrid approaches also have a significant contribution with 10 contributions. While the machine learning and rule-based approaches have fewer contributions, showing growing preference for advanced approaches.

Answer to research question 2

Research question 2: What algorithms were previously used for sentiment analysis in low-resource languages?

Table 4: Algorithms for sentiment analysis in low-resource languages

Authors	Algorithms
Roy (2024)	BERT, ANN
Hashem et al. (2024)	BiLSTM
Yu et al. (2024)	BUQRNN
Onan & Balbal (2024)	BERT-Large-ITPT XLNet, BERTurk, DistilBERTurk, ULM-Fit, SSL, ELECTRA, and ConvBERTurk
Husain et al. (2024)	ARBERT

Saroj et al. (2024)	BiLSTM
Wang & Xiong (2024)	BERT-BiLSTM
Cosme & de-Leon (2024)	XLM-RoBERTa
Alqurafi & Alsanoosy (2024)	LR, SVM, DT, and LSTM
Mohammed & Prasad (2023)	BERT
Yusup et al. (2023)	A hybrid of TextCNN, RNN, RCNN, and LSTM
Ghasemi & Momtazi (2023)	XLM-RoBERTa
Vo et al. (2023)	LSTM, BERT, DistilBERT, BERT-RCNN and PhoBERT
Rasool et al. (2023)	ABiLSTM
Raychawdhary et al. (2023)	AfriBERTa
Hancharova et al. (2023)	SVM
Akrah & Pedersen (2023)	TwiBERT
Garcia-Diaz et al. (2023a)	RoBERTa
Salahudeen et al. (2023)	AfriBERTa-Large, BERT, mBERT,Afro-xlmr-Large, and Arabic-CamelBERT
Ramanathan et al. (2023)	RoBERTa
García-Díaz et al. (2023b)	Language-Independent Linguistics
Prakash & Vijay (2023)	MSDLA
Abdirahman et al. (2023)	DT Classifier, RF Classifier, XGB, CNN, and LSTM
Mujahid et al. (2023)	RoBERTa-BERT
Ullah et al. (2022)	CNN-LSTM
Altaf et al. (2022)	RNN, LSTM, and GRU
Shanmugavadivel et al. (2022)	LR, CNN, Bi-LSTM, BERT, RoBERTa, and Adapter-BERT
Das & Singh (2022)	Multimodal Framework
Mamta et al. (2022)	BERT
Khan et al. (2022)	SVM, NB, Adaboost, MLP, LR, RF, CNN, LSTM, Bi-LSTM, GRU, Bi-GRU, mBERT
Bensoltane & Zaki (2022)	BERT
Garg, K., & Lobiyal, D. K. (2021)	neuro-fuzzy
Tho et al. (2021)	Sentence-BERT
Kastrati et al. (2021)	BiLSTM
Bhowmick & Jana (2021)	XLM-RoBERTa
Khamphakdee & Seresangtakul (2021)	SVM
Divate (2021)	LSTM
Bharti et al. (2020)	TNWS, TIWS, and RiFoQ
Xu et al. (2020)	DomBERT
Singh et al. (2020)	BiLSTM
Aliramezani et al. (2020)	RNN

Table 4 shows the previous low-resource language sentiment analysis algorithms in the past five years. The table shows that BERT-based algorithms such as RoBERTa and XLM-RoBERTa are mostly used due to their adaptability abilities. Deep learning-based models such as LSTM, BiLSTM, and CNN are also widely used, alongside hybrid approaches like BUQRNN and neuro-fuzzy systems. Traditional methods such as SVM, RF, and LR have remained relevant over the past 5 years.

Answer to research question 3.
Research question 3: How well have the previously used algorithms for sentiment analysis in low-resource languages performed?
Table 5: Performance Evaluation of Sentiment Analysis Algorithms for Low-Resource Languages

Algorithm	Authors	Accuracy/F1-Score	Notes
BERT	Roy (2024), Mohammed & Prasad (2023)	98%, 72% accuracy	High adaptability, best for English and other widely spoken languages
RoBERTa	Mujahid et al. (2023), Ramanathan et al. (2023)	99.94% F1-score (train), 80.32% F1-score (test)	High performance on specific languages, especially when fine-tuned
BUQRNN (Hybrid)	Yu et al. (2024)	99.3% accuracy	Hybrid model combining BiLSTM and QNN for improved accuracy
BiLSTM	Saroj et al. (2024), Hashem et al. (2024)	90.9% – 94.6% accuracy	Good for sequential data, effective for languages with complex structures
XLM-RoBERTa	Cosme & de-Leon (2024), Bhowmick & Jana (2021)	84% F1-score	Performs well with multilingual datasets, not always

			optimal for low-resource languages
LSTM	Abdirahman et al. (2023), Altaf et al. (2022)	93.3% – 94.6% accuracy	Strong for sequential tasks, but requires large datasets for optimal results
SVM	Hancharova et al. (2023), Khamphakdee & Seresangtakul (2021)	64% – 91.3% accuracy	Traditional machine learning, still relevant for simple sentiment tasks
Hybrid Approaches (e.g., TextCNN, RNN, RCNN)	Yusup et al. (2023)	82.17% accuracy	Combines different deep learning techniques for improved performance

Table 5 displays the performance evaluation of the algorithms used for sentiment analysis in low-resource languages. From the table, we observed that the Transformer-Based Approaches with BERT related model always outperform all other compared approaches, achieving accuracy up to 99.94% and F1-score up to 99%. This shows the model's high adaptability to low-resource languages when fine-tuned for specific datasets. However, the effectiveness of these models may be limited for languages with insufficient training data. Also, it is observed that Deep Learning Models like LSTM and BiLSTM also perform well in most languages, with accuracies of up to 94.6%. The Hybrid Approaches like the BUQRNN model were able to achieve an accuracy of up to 99.3%, this shows the potential of combining different deep learning models for high performance in low-resource languages. The traditional machine learning models like the SVM, and LR had lower performance compared to other approaches. With accuracy ranging from 64% to 91.3%, this indicates that these models are less effective as compared to deep learning and transformer-based models, but they remain relevant for simpler tasks and specific applications. Generally, we observed low performance of these algorithms in some specific datasets and underrepresented languages. The low performance that was observed indicates the challenges in adapting these models to diverse low-resource language contexts.

Discussion of Results

The result from this study revealed the transformer-based approaches, particularly BERT-related approaches as the most used and most effective approach for low-resource language sentiment analysis due to their performance and adaptability. The deep learning approach has also demonstrated good performance with high accuracies and precisions, the deep learning-based models like LSTM and BiLSTM show significant usage and high accuracies to reflect their robustness. Also, it was observed that the Hybrid approaches that combine BUQRNN and neuro-fuzzy systems algorithms have performed well with notable accuracy improvement. However, machine learning-based approaches like SVM and LR are less prevalent in this context but remain relevant for specific applications. Performance analysis of the recently used algorithms shows that the transformer-based algorithms, like the BERT achieved accuracy up to 98% and RoBERTa achieved F1 scores up to 99.94% F1-score, making them outperform others consistently, establishing their superiority. The deep learning-based models particularly the LSTM and the BiLSTM perform very well achieving an accuracy of up to 94.6%. Hybrid methods like BUQRNN achieved up to 99.3% accuracy, showcasing their potential to improve sentiment analysis outcomes.

Conclusion

In this study, we reviewed various approaches, and algorithms, as well as their performance in sentiment analysis of low-resource languages. Thus, we were able to provide comprehensive insight into the advancement of the field. Based on the study inclusion and exclusion criteria, we were able to collect 42 research papers from four different academic databases which include Scopus, IEEE, Springer Link, and Web of Science. After a thorough review of this research paper, we identified five primary approaches for sentiment analysis over the past 5 years, these approaches include machine learning, deep learning, transformer-based, rule-based, and hybrid approaches. This study observed that the transformer-based approach is mostly used for sentiment analysis of low-resource languages, this is due to their adaptability and robustness, while the deep learning and hybrid approaches are growing fast in popularity among researchers. Another observation in this research is that the BERT-based algorithms have the best performance in low-resource language sentiment analysis, showcasing superior adaptability and performance when compared to the classical machine learning models like SVM and LR. Despite the different approaches used in sentiment analysis of low-resource languages, our study shows that the transformer-based algorithms constantly achieved higher accuracy and efficiency than the compared approaches, this highlights their potential for future research in this domain.

Recommendations

This research recommends that future research focus on developing annotated datasets for low-resource languages by collaborating with researchers and linguistic experts. Also, this research suggests that improvement should be made to existing algorithms and new models should be tailored to unique linguistic characteristics of low-resource languages.

Conflict of Interests

The authors declare no conflict of interest.

References

Abdirahman, A. A., Hashi, A. O., Dahir, U. M., Elmi, M. A., & Rodriguez, O. E. R. (2023). Comparative Analysis of Machine Learning and Deep Learning Models for Sentiment Analysis in Somali Language. *SSRG International Journal of Electrical and Electronics Engineering*, *10*(7). https://doi.org/10.14445/23488379/IJEEE-V10I7P104

Akrah, S., & Pedersen, T. (2023). DuluthNLP at SemEval-2023 Task 12: AfriSenti-SemEval: Sentiment Analysis for Low-resource African Languages using Twitter Dataset. *17th International Workshop on Semantic Evaluation, SemEval 2023 - Proceedings of the Workshop*.

Aliramezani, M., Doostmohammadi, E., Bokaei, M. H., & Sameti, H. (2020). Persian Sentiment Analysis without Training Data Using Cross-Lingual Word Embeddings. *2020 10th International Symposium on Telecommunications: Smart Communications for a Better Life, IST 2020*. https://doi.org/10.1109/IST50524.2020.9345882

Alqurafi, A., & Alsanoosy, T. (2024). Measuring Customers' Satisfaction Using Sentiment Analysis: Model and Tool. *Journal of Computer Science*, *20*(4). https://doi.org/10.3844/jcssp.2024.419.430

Altaf, A., Anwar, M. W., Jamal, M. H., Hassan, S., Bajwa, U. I., Choi, G. S., & Ashraf, I. (2022). Deep Learning Based Cross-Domain Sentiment Classification for Urdu Language. *IEEE Access*, *10*. https://doi.org/10.1109/ACCESS.2022.3208164

Baniata, L. H., & Kang, S. (2024). Switch-Transformer Sentiment Analysis Model for Arabic Dialects That Utilizes a Mixture of Experts Mechanism. *Mathematics*, *12*(2). https://doi.org/10.3390/math12020242

Bensoltane, R., & Zaki, T. (2022). Towards Arabic aspect-Based sentiment analysis: a transfer learning-based approach. *Social Network Analysis and Mining*, *12*(1). https://doi.org/10.1007/s13278-021-00794-4

Bharti, S. K., Naidu, R., & Babu, K. S. (2020). Hyperbolic Feature-based Sarcasm Detection in Telugu Conversation Sentences. *Journal of Intelligent Systems*, *30*(1). https://doi.org/10.1515/jisys-2018-0475

Bhowmick, A., & Jana, A. (2021). Sentiment Analysis For {B}engali Using Transformer Based Models. *Proceedings of the 18th International Conference on Natural Language Processing (ICON), 2020*.

Bu, K., Liu, Y., & Ju, X. (2024). Efficient utilization of pre-Trained models: A review of sentiment analysis via prompt learning. *Knowledge-Based Systems*, *283*. https://doi.org/10.1016/j.knosys.2023.111148

Cofino, C. L., Escorial, R. B., & Enquilino, D. L. B. (2024). A Literature Review on Natural Language Processing (NLP) in Aiding Industry to Progress. In *International Journal of Engineering Trends and Technology* (Vol. 72, Issue 2). https://doi.org/10.14445/22315381/IJETT-V72I2P105

Cosme, C. J., & de Leon, M. M. (2024). Sentiment Analysis Of Code-Switched Filipino-English Product and Service Reviews Using Transformers-Based Large Language Models. *Lecture Notes in Networks and Systems*, *834*. https://doi.org/10.1007/978-981-99-8349-0_11

Cui, J., Wang, Z., Ho, S. B., & Cambria, E. (2023). Survey On sentiment analysis: evolution of research methods and topics. *Artificial Intelligence Review*, *56*(8). https://doi.org/10.1007/s10462-022-10386-z

Das, R., & Singh, T. D. (2022). A multi-stage multimodal Framework for sentiment analysis of Assamese in a low resource setting. *Expert Systems with Applications*, *204*. https://doi.org/10.1016/j.eswa.2022.117575

Divate, M. S. (2021). Sentiment analysis of Marathi news Using LSTM. *International Journal of Information Technology (Singapore)*, *13*(5). https://doi.org/10.1007/s41870-021-00702-1

García-Díaz, J. A., Caparros-Laiz, C., Almela, Á., Alcaráz-Mármol, G., Marín-Pérez, M. J., & Valencia-García, R. (2023b). UMUTeam at SemEval-2023 Task 12: Ensemble Learning of LLMs Applied to Sentiment Analysis for Low-resource African Languages. *17th International Workshop on Semantic Evaluation, SemEval 2023 - Proceedings of the Workshop*. https://doi.org/10.18653/v1/2023.semeval-1.38

Garcia-Diaz, J. A., Garcia-Sanchez, F., & Valencia-Garcia, R. (2023a). Smart Analysis of Economics Sentiment in Spanish Based on Linguistic Features and Transformers. *IEEE Access*, *11*. https://doi.org/10.1109/ACCESS.2023.3244065

Garg, K., & Lobiyal, D. K. (2021). KL-NF technique for Sentiment classification. *Multimedia Tools and Applications*, *80*(13). https://doi.org/10.1007/s11042-021-10559-y

Ghasemi, R., & Momtazi, S. (2023). How a Deep Contextualized Representation and Attention Mechanism Justifies Explainable Cross-Lingual Sentiment Analysis. *ACM Transactions on Asian and Low-Resource Language Information*

Processing, *22*(11). https://doi.org/10.1145/3626094

Hancharova, A., Wang, J., & Kumar, M. (2023). Team ISCL_WINTER at SemEval-2023 Task 12:AfriSenti-SemEval: Sentiment Analysis for Low-resource African Languages using Twitter Dataset. *17th International Workshop on Semantic Evaluation, SemEval 2023 - Proceedings of the Workshop*. https://doi.org/10.18653/v1/2023.semeval-1.149

Hashem, S. A., El-Bendary, N., & Elhariri, E. (2024). Autoregressive Feature Extraction with Topic Modeling for Aspect-based Sentiment Analysis of Arabic as a Low-resource Language. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *23*(2). https://doi.org/10.1145/3638050

Husain, F., Alostad, H., & Omar, H. (2024). Bridging the Kuwaiti Dialect Gap in Natural Language Processing. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3364367

Kastrati, Z., Ahmedi, L., Kurti, A., Kadriu, F., Murtezaj, D., & Gashi, F. (2021). A deep learning sentiment analyzer for social media comments in low-resource languages. *Electronics (Switzerland)*, *10*(10). https://doi.org/10.3390/electronics10101133

Khamphakdee, N., & Seresangtakul, P. (2021). Sentiment Analysis of the Thai language in the hotel domain using machine learning algorithms. *Acta Informatica Pragensia*, *10*(2). https://doi.org/10.18267/j.aip.155

Khan, L., Amjad, A., Ashraf, N., & Chang, H. T. (2022). Multi-class sentiment analysis of Urdu text using multilingual BERT. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-09381-9

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering - A systematic literature review. In *Information and Software Technology* (Vol. 51, Issue 1). https://doi.org/10.1016/j.infsof.2008.09.009

Li, C., Zhang, Y., Weng, Y., Wang, B., & Li, Z. (2023). Natural Language Processing Applications for Computer-Aided Diagnosis in Oncology. In *Diagnostics* (Vol. 13, Issue 2). https://doi.org/10.3390/diagnostics13020286

Mamta, Ekbal, A., & Bhattacharyya, P. (2022). Exploring Multi-lingual, Multi-task, and Adversarial Learning for Low-resource Sentiment Analysis. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *21*(5). https://doi.org/10.1145/3514498

Mohammed, I., & Prasad, R. (2023). Building lexicon-based Sentiment analysis model for low-resource languages. *MethodsX*, *11*. https://doi.org/10.1016/j.mex.2023.102460

Mujahid, M., Kanwal, K., Rustam, F., Aljadani, W., & Ashraf, I. (2023). Arabic ChatGPT Tweets Classification Using RoBERTa and BERT Ensemble Model. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(8). https://doi.org/10.1145/3605889

Onan, A., & Balbal, K. F. (2024). Improving Turkish Text Sentiment Classification Through Task-Specific and Universal Transformations: An Ensemble Data Augmentation Approach. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3349971

Prakash, J. v., & Vijay, A. A. S. (2023). Cross-lingual Sentiment Analysis of Tamil Language Using a Multi-stage Deep Learning Architecture. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(12). https://doi.org/10.1145/3631391

Ramanathan, N., Sivanaiah, R., Suseelan, A. D., & Thanagathai, M. T. N. (2023). TechSSN at SemEval-2023 Task 12: Monolingual Sentiment Classification in Hausa Tweets. *17th International Workshop on Semantic Evaluation, SemEval 2023 - Proceedings of the Workshop*. https://doi.org/10.18653/v1/2023.semeval-1.165

Raychawdhary, N., Das, A., Dozier, G., & Seals, C. D. (2023). Seals_Lab at SemEval-2023 Task 12: Sentiment Analysis for Low-resource African Languages, Hausa and Igbo. *17th International Workshop on Semantic Evaluation, SemEval 2023 -Proceedings of the Workshop*. https://doi.org/10.18653/v1/2023.semeval-1.208

Roy, P. K. (2024). Deep Ensemble Network for Sentiment Analysis in Bi-lingual Low-Resource Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *23*(1). https://doi.org/10.1145/3600229

Salahudeen, S. A., Lawan, F. I., Wali, A. M., Imam, A. A., Shuaibu, A. R., Yusuf, A., Rabiu, N. B., Bello, M., Adamu, S. U., Aliyu, S. M., Gadanya, M. S., Muaz, S. A., Ahmad, M. S., Abdullahi, A., & Jamoh, A. Y. (2023). HausaNLP at SemEval-2023 Task 12: Leveraging African Low Resource TweetData for Sentiment Analysis. *17th International Workshop on Semantic Evaluation, SemEval 2023 - Proceedings of the Workshop*. https://doi.org/10.18653/v1/2023.semeval-1.6

Saroj, A., Thakur, A., & Pal, S. (2024). Sentiment analysis on Hindi tweets during COVID-19 pandemic. *Computational Intelligence*, *40*(1). https://doi.org/10.1111/coin.12622

Shanmugavadivel, K., Sathishkumar, V. E., Raja, S., Lingaiah, T. B., Neelakandan, S., & Subramanian, M. (2022). Deep learning-based sentiment analysis and offensive language identification on multilingual code-mixed data. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-26092-3

Singh, O. M., Timilsina, S., Bal, B. K., & Joshi, A. (2020).

Aspect-Based Abusive Sentiment Detection in Nepali Social Media Texts. *Proceedings of the 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2020*. https://doi.org/10.1109/ASONAM49781.2020.9381292

Tho, C., Heryadi, Y., Kartowisastro, I. H., & Budiharto, W. (2021). A Comparison of Lexicon-based and Transformer-based Sentiment Analysis on Code-mixed of Low-Resource Languages. *Proceedings of 2021 1st International Conference on Computer Science and Artificial Intelligence, ICCSAI 2021*. https://doi.org/10.1109/ICCSAI53272.2021.9609781

Ullah, F., Chen, X., Shah, S. B. H., Mahfoudh, S., Hassan, M. A., & Saeed, N. (2022). A Novel Approach for Emotion Detection and Sentiment Analysis for Low Resource Urdu Language Based on CNN-LSTM. *Electronics (Switzerland)*, *11*(24). https://doi.org/10.3390/electronics11244096

Vo, Q. T., Tran, P., & Tran, T. (2023). Sentiment analysis for A low-resource language: A study on a Vietnamese University. *ICIC Express Letters, Part B: Applications*, *14*(11). https://doi.org/10.24507/icicelb.14.11.1115

Wang, P., & Xiong, X. (2024). A BERT-ABiLSTM Hybrid Model-Based Sentiment Analysis Method for Book Review. *Journal of Circuits, Systems, and Computers*, *33*(2). https://doi.org/10.1142/S0218126624500397

Xu, H., Liu, B., Shu, L., & Yu, P. S. (2020). DomBERT: Domain-oriented language model for aspect-based sentiment analysis. *Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020*. https://doi.org/10.18653/v1/2020.findings-emnlp.156

Yu, W., Yin, L., Zhang, C., Chen, Y., & Liu, A. X. (2024). Application of Quantum Recurrent Neural Network in Low-Resource Language Text Classification. *IEEE Transactions on Quantum Engineering*, *5*. https://doi.org/10.1109/TQE.2024.3373903

Yusup, A., Chen, D., Ge, Y., Mao, H., & Wang, N. (2023). Resource Construction and Ensemble Learning Based Sentiment Analysis for the Low-Resource Language Uyghur. *Journal of Internet Technology*, *24*(4). https://doi.org/10.53106/160792642023072404018