IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Systematic Review Of Aspect Based Sentiment Analysis For Marathi Language

[1]Dr. Dhiraj Amin , [2]Anushka Raut , [3]Piyul Worlikar , [4]Kanchan Patil , [5]Vedika Jadhav
[1]Professor , [2]Student , [3]Student , [4]Student , [5]Student
[1]Department of Computer Engineering ,
[1]Pillai College of Engineering ,New Panvel, India

Abstract: Sentiment analysis is the process of analyzing online conversations like tweets, blog posts, or comments to understand people's opinions about specific topics or services. It categorizes these opinions as positive, negative, or neutral, helping businesses assess customer sentiment regarding their products.Aspect-Based Sentiment Analysis (ABSA) takes this a step further by identifying specific features or aspects of a product mentioned in customer reviews. It also determines whether the sentiment about each aspect is positive, negative, or neutral. This detailed feedback helps businesses focus on improving specific product features based on customer opinions, making their products more appealing and successful.The proposed method for ABSA uses BERT, a powerful pre-trained language model, to better understand the context in reviews. By fine-tuning BERT on data specific to a particular domain and using techniques that address multiple tasks at once (like extracting aspects and classifying sentiment), this approach improves the accuracy of ABSA. It is especially effective at handling complex language and understanding relationships between different aspects in reviews.

***Index Terms* - Aspect-Based Sentiment Analysis , Sentiment Analysis , Marathi Language Processing , Natural Language Processing (NLP) , BERT.**

I. INTRODUCTION

Sentiment analysis, also known as opinion mining, is the process of understanding the feelings or opinions expressed in a text, such as a review, comment, or social media post. Its goal is to determine whether the sentiment is positive, negative, or neutral.

Here are three common types of sentiment analysis:

1. **Fine-grained Sentiment Analysis**

   This type measures emotions on a scale, like a rating from 0 to 100, similar to how star ratings are used to gauge customer satisfaction. It captures not just the sentiment but also the strength or degree of emotion conveyed in the text.

2. **Aspect-Based Sentiment Analysis (ABSA):**
ABSA focuses on specific aspects or feature Of a product, service, or experience. For instance, a travel app might use ABSA to analyze customer feedback about the user interface or a chatbot's performance. This method helps businesses understand which aspects of their offerings are working well and which need improvement.

3. **Emotion Detection**

This type digs deeper to identify the mindset of the writer, such as frustration, happiness, or indifference. Unlike other methods that categorize text as either positive, negative, or neutral, emotion detection provides a deeper and more detailed understanding of the writer's psychological state and intentions.

ABSA is a machine learning technique that identifies specific aspects in a text and assigns sentiment to each. It provides a deeper level of analysis compared to general sentiment analysis.
For example:

- **Document-based sentiment analysis** gives an overall sentiment for the entire text.
- **Topic-based sentiment analysis** identifies sentiment for a specific topic within the text.
- **ABSA**, however, breaks the text down into finer details, analyzing sentiments about individual aspects like product features or services.

Let's consider the sentence:
ष्एपेटाइजर ठीक होतेए पेये इतके खास नव्हते आणि वातावरण खूपच खराब होतेण्

"The appetizers were fine, the drinks weren't that great, and the atmosphere was terrible."
Here's how ABSA would analyze it:

- For food (खाणे), the sentiment is neutral.
- For drinks (पेये), the sentiment leans negative.
- For atmosphere (वातावरण), the sentiment is strongly negative.

By breaking down the text, ABSA helps businesses understand customer opinions about specific aspects, making it easier to identify areas of improvement.

II. RESEARCH METHODOLOGY

This section includes a system of guidelines for designing and analysing the studies of aspect based sentiment analysis of various asian languages. The paper titled "Sentiment Analysis of Marathi Language," authored by Sujata Deshmukh, Nileema Patil, Surabhi Rotiwar, and Jason Nunes, was published in June 2017. The authors address the difficulties of sentiment analysis in the Marathi language, recognizing the increasing significance of regional languages in India's online environment. Their proposed system employs a corpus-based approach, creating an up-to-date corpus of Marathi keywords with individual polarities, utilizing WordNet. The algorithm calculates cumulative polarity, categorizing sentences as either positive, negative, or neutral. The research methodology is conducted by applying the following review strategies: (1) designing research questions;(2) searching related aspect based sentiment analysis or sentiment analysis (3) applying predetermined inclusion and exclusion criteria; and (4) applying quality assessment criteria.

A. RESEARCH QUESTION

This study aims to answer the following research questions to highlight critical practical aspects of aspect based sentiment analysis or sentiment analysis in marathi language. The four research questions addressed in this literature review are as follows:

RQ1: Which techniques and features have been used for ABSA in Marathi?

This RQ1 addresses various techniques that have been used for ABSA in other languages, and many of these methods can be adapted for Marathi. The methods include machine learning algorithms like Support Vector Machine (SVM), Naïve Bayes (NB), and Random Forest (RF), as well as deep learning approaches like Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and BERT-based models. Some studies also experimented with hybrid models, combining traditional machine learning with deep learning techniques or optimizing models using genetic algorithms (GA). For Marathi ABSA, similar techniques could be applied, but special attention must be given to Marathi's linguistic characteristics, such as its morphology and syntactic structure,to ensure that the models effectively capture sentiment related to specific aspects.

RQ2: What challenges exist in ABSA for Marathi?

In ABSA for Marathi, several challenges are similar to those in other low-resource languages, such as limited annotated datasets and the complexity of linguistic features. Marathi has a rich morphology and flexible word order, which can make sentiment classification challenging, especially when trying to extract aspect-based sentiment. Moreover, the lack of large, high-quality Marathi datasets, similar to the challenges seen in Bengali or Telugu, hampers model training and evaluation. Additionally, cultural aspects such as idiomatic expressions, region-specific sentiment expressions, and domain-specific variations (e.g., the way food or ambiance is discussed in reviews) can add further complexity. The creation of resources for Marathi ABSA will need to address these difficulties by creating more extensive, annotated corpora and fine-tuning existing models for Marathi's unique linguistic and cultural traits.

RQ3: What datasets are available for Marathi ABSA?

Currently, there is limited availability of publicly available datasets specifically for ABSA in Marathi, similar to the situation in other South Asian languages like Bengali or Nepali. However, there are ongoing efforts to create datasets for related tasks, such as sentiment classification and aspect extraction. For example, datasets for sentiment analysis in Marathi exist, but they may not be specifically annotated for aspect-based sentiment classification. To improve ABSA in Marathi, datasets need to be developed with a focus on aspect term extraction, sentiment polarity, and domain-specific data (like reviews for restaurants, products, or services). These datasets must cover a range of topics, be annotated in detail, and reflect the varied linguistic styles of Marathi speakers to ensure robust ABSA models.

RQ4: What is the proposed workflow for ABSA in Marathi for future research? Developing a standardized framework for future ABSA research.

The proposed workflow for Marathi ABSA is starting with the collection of diverse Marathi texts, including social media, reviews, and news, while focusing on aspect-based analysis. Preprocessing has addressed the unique linguistic features of Marathi, including tokenization, stopword removal, and part-of-speech tagging. For model selection, approaches like CNN, LSTM, BiLSTM, or BERT can be used, along with hybrid models to enhance performance. The workflow should ensure accurate aspect detection and sentiment classification, evaluated through metrics like accuracy and F1-score. Continuous improvement should involve retraining models with new data and fine-tuning them for better results. This approach adapts existing techniques to Marathi's linguistic challenges.

A. DATA SOURCES AND SEARCH STRATEGY

In this work, we referred to various research papers to find related articles through electronic databases. We selected five electronic sources to gather our references. Through the electronic sources, we investigated all available materials pertaining to the objectives of this systematic literature review. Search strings (keywords) were developed to collect related research papers responding to address the research questions. The search strings were created using critical terms within the topic field and the objective of the pape

B. INCLUSION AND EXCLUSION CRITERIA

This section discusses the inclusion and exclusion criteria used in our literature study. Meta-data and abstracts of papers were reviewed to determine which studies should be included in the review and removed irrelevant articles. The following criteria were applied for inclusion: (I1) Studies published between 2020 and 2024; (I2) full-text papers; (I3) papers written for Asian languages; (I4) papers related to Aspect Based Sentiment Analysis and Sentiment Analysis. We excluded those articles that did not satisfy the criteria for inclusion from the study. Also, any publications that did not match any of the excluded criteria were excluded.

The inclusion and exclusion criteria for this study are presented in Table 3. The following are the exclusion criteria to eliminate irrelevant papers: (E1) papers not written for marathi or asian languages; (E2) papers that do not focus on solving the problem of sentiment analysis; (E3) papers that do not discuss ABSA or Sentiment Analysis.

Table 1. Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
• Papers published within the period of 2020 to 2024. • Full-text papers • Paper written for Marathi or asian languages • Papers related to ABSA or Sentiment Analysis	• papers not written for marathi or asian languages • papers that do not focus on solving the problem of sentiment analysis • papers that do not discuss ABSA or Sentiment Analysis

C. QUALITY ASSESSMENT (QA)

The QA is applied in this systematic review of the literature to assess the strength of the selected studies. The QA was developed using tools like a checklist of all aspects or queries required to be applied to each study. The following questions were developed as the QA criteria for each study:

1. (C1) Does the paper describe the marathi dataset clearly?
2. (C2) Are the techniques clearly explained in the paper?
3. (C3) Does the research paper explain the challenges in ABSA for marathi language?
4. (C4) Are the findings clearly presented in the paper?

D. STUDY SELECTION PROCESS

This section explains the selection process to determine relevant studies that meet all the research questions. The study selection task is done in four phases: identification, screening, eligibility, and included studies. In the identification stage, we searched the literature from five electronic databases using predefined keywords and obtained 100 research papers. Firstly, we screened the retrieved research by removing duplicate papers. A total of 20 papers were removed in the screening stage. We applied inclusion and exclusion criteria to 80 articles, and a total of 10 studies were eliminated. After that, the rest of the 70 articles were assessed using the five quality criteria, and we excluded 55 papers in this stage. Finally, a collection of 15 research papers (referred to as selected studies) was included in this literature review, 4 papers (26%) from journals and 11 papers (73%) from conferences.

III.RESULT AND DISCUSSION

This section summarizes the key findings from primary studies on Marathi text. The discussion is divided into three parts, each addressing a specific research question (RQ).The first subsection addresses RQ1 regarding existing techniques for aspect-based sentiment analysis in Marathi text. In the second subsection, we present our findings concerning RQ2, which investigates the challenges of performing aspect-based sentiment analysis on Marathi text. Subsequently, we provide our findings regarding dataset availability and quality criteria for evaluating aspect-based sentiment analysis tasks. For this, we provide a framework in response to RQ3, which is explained in the last subsection.

(RQ1) WHICH APPROACHES HAVE BEEN USED FOR ABSA OF MARATHI LANGUAGE?

A. Machine Learning Approaches

Traditional machine learning methods have been widely used for sentiment analysis, relying on manual feature extraction and standard algorithms. Common techniques for extracting features include TF-IDF, Bag of Words (BoW), and n-grams. For classification, algorithms like Support Vector Machines (SVM), Naïve Bayes (NB), Logistic Regression (LR), and Random Forest (RF) are commonly applied.

For instance, in Sentiment Analysis in Marathi Language[1], SVM and NB were utilized to classify sentiment within Marathi text datasets. Similarly, Exploring Sentiment Analysis in Kannada Language: A Comprehensive Study on COVID-19 Data employed machine learning models[12] such as XGBoost, LR, and RF to analyze COVID-19-related Kannada tweets. Another notable work, Aspect-Based Sentiment Analysis of Bangla[10] Comments on Entertainment Domain, incorporated classifiers such as Support Vector Classifier (SVC), RF, and LR to evaluate Bangla comments in the entertainment domain.In more specific applications, Named Entity Recognition and Aspect-Based Sentiment Analysis employed a combination of tools including Tweepy, NLTK, VADER, and Orange to extract entities and classify sentiment. Similarly, Facebook for Sentiment Analysis: Baseline Models to Predict Facebook Reactions of Sinhala Posts developed multiple models, such as Core Reaction Set, All Reaction Set, and Star Rating models, to predict reactions to Sinhala Facebook posts.

B. Deep Learning Approaches

Deep learning techniques have significantly advanced sentiment analysis by using models that can understand complex language patterns and context. Architectures like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and transformer-based models such as BERT have been widely applied. For example, Towards Emotion-Aware Multimodal Marathi Sentiment Analysis[7] used CNN, Bi-LSTM, MahaRoBERTa, and IndicBERT to analyze Marathi sentiment across different types of data. Similarly, Sentiment Analysis for Sinhala employed advanced models like hierarchical attention hybrid networks and capsule networks to analyze Sinhala text. In Nepali COVID-19 Tweets, LSTM models were combined with SVM and NB to analyze tweets. A comparative study on Indonesian Hotel Reviews tested several models including CNN, LSTM, GRU, Bi-LSTM, and combinations like CNN-LSTM and CNN-BiLSTM for aspect-based sentiment analysis. Indonesian Tourist Attraction Reviews highlighted the effectiveness of Bi-LSTM in identifying aspect-based sentiments. Additionally, a Multi-Task Learning Model for Arabic Sentiment Analysis used Bi-LSTM-CRF for aspect extraction and BERT for sentiment classification, demonstrating the power of deep learning in understanding sentiment at a granular level.

C. Hybrid Approaches

Hybrid approaches in sentiment analysis combine multiple models and techniques to improve the accuracy and reliability of results. These methods often blend CNNs with sequence models like LSTM or GRU, along with transformer models such as BERT. For example, the study Sentiment Analysis of marathi news using LSTM[3] showed the effectiveness of hybrid models like CNN-GRU, CNN-LSTM, and stacked Bi-LSTMs in analyzing large Sinhala datasets. Similarly, ABSA of Indonesian Customer Reviews Using IndoBERT [16]combined the IndoBERT model with CNN and traditional classifiers like XGBoost, RF, and NB to perform aspect-based sentiment analysis. In another study, A Dataset for Multi-modal Sentiment Analysis in Tamil and Malayalam [17] used hybrid models to integrate both text and multimodal features for sentiment classification in these languages. Additionally, An Approach for Analyzing Tamil Movie Reviews Using Tamil Tweets [18] combined TF-IDF with advanced techniques like Domain-Specific Ontology (DSO) and Contextual Sentiment Strength Analysis (CSSA) to analyze aspect-based sentiment in Tamil tweets. These examples highlight how hybrid approaches improve the effectiveness of sentiment analysis across different languages and context.

Figure 1 : Taxonomy of Approaches for ABSA for marathi language

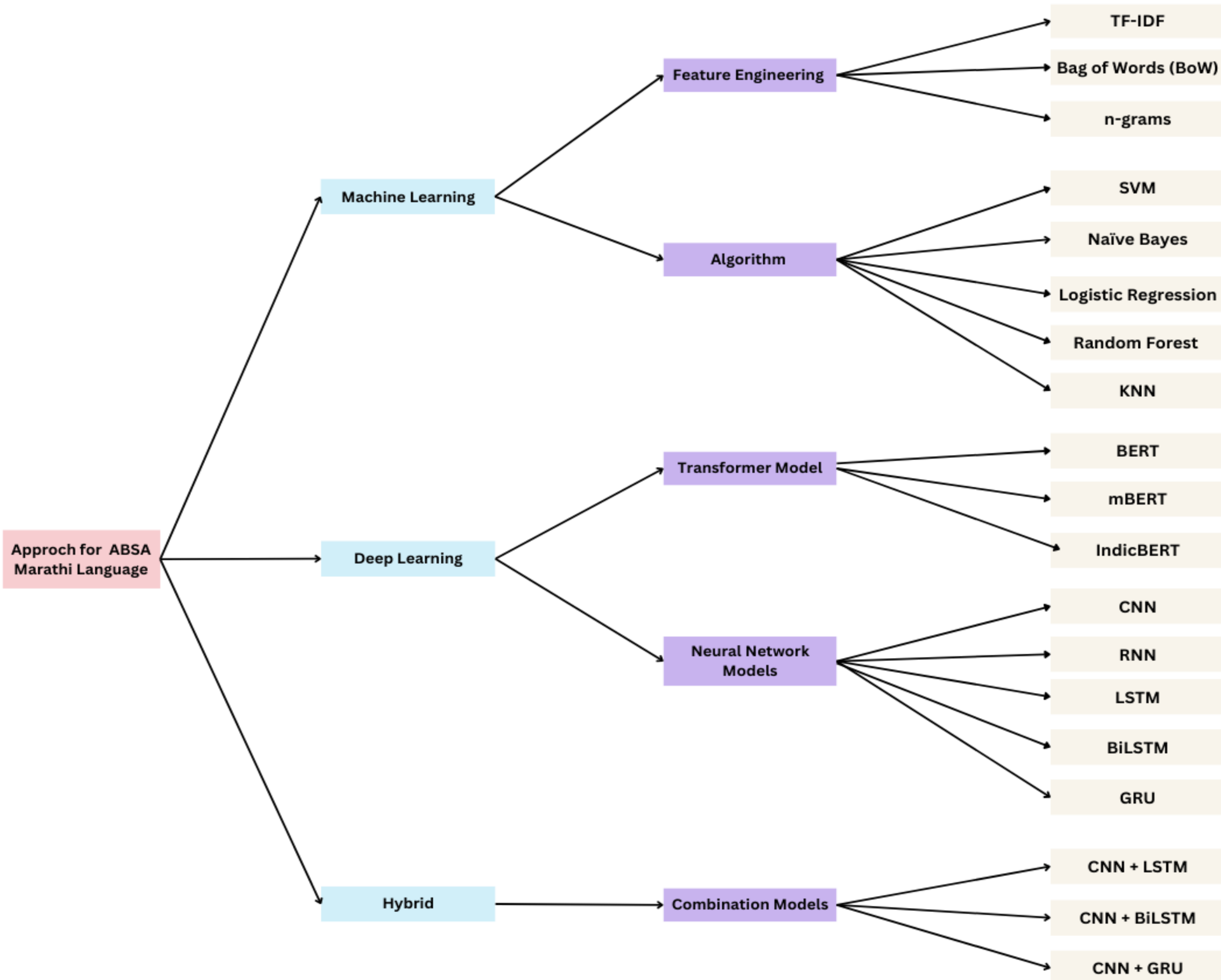

Table 2 : Language , Applied Techniques & Reported Performance . This table is sorted by published year from 2017 to 2024

References	Title	Year	Language	Applied Techniques
[1]	SENTIMENT ANALYSIS OF MARATHI LANGUAGE	2017	Marathi	CORPUS GENERATION FROM ENGLISH SENTI WORD NET
[2]	SENTIMENT ANALYSIS OF MIXED CODE FOR THE TRANSLITERATED HINDI AND MARATHI TEXTS	2018	Hindi & Marathi	mBERT
[4]	A Marathi Tweet-based Sentiment Analysis	2021	Marathi	CNN, BiLSTM+GlobalMaxPool,

	Dataset			ULMFiT , BERT
[3]	Sentiment analysis of Marathi news using LSTM	2021	Marathi	Support vector machine (SVM), Naıve Bayes, LSTM,CNN, Combine CNN-LSTM
[14]	A Review on Recognition of Sentiment Analysis of Marathi Tweets using Machine Learning Concept	2021	Marathi	Support vector machine (SVM), Naıve Bayes, LSTM,CNN,
[5]	Lexical Resource Creation and Evaluation: Sentiment Analysis in Marathi	2022	Marathi	Logistic Regression Stochastic Gradient Descent (SGD) Support Vector Machine (SVM) Nearest Neighbour Neural Network Decision Tree (DT) Naïve Bayes (NB) Ensemble-based Model
[6]	A Multi-domain Marathi Sentiment Analysis Dataset and Transformer Models	2023	Marathi	mBERT , indicBERT , MuRIL , mahaBERT
[7]	Towards Emotion-aware Multimodal Marathi Sentiment Analysis	2023	Marathi	CNN , Bi-LSTM MURIL ,XLM,MahaRoBERTa ,MarathiAlBERT, MarathiSentiment and IndicBERT
[15]	Sentence Annotation for Aspect-oriented Sentiment Analysis: A Lexicon based Approach with Marathi Movie Reviews	2024	Marathi	Naïve bayes classifier (NBC) Decision tree classifier (DTC)

B. (RQ2) WHAT ARE THE CHALLENGES IN ABSA FOR MARATHI LANGUAGE ?

A. Limited Availability of Datasets

Marathi, being a low-resource language, suffers from a significant lack of annotated datasets required for training and evaluating ABSA models. This scarcity extends to domain-specific corpora, such as product reviews or customer feedback, which are crucial for building models capable of performing well across diverse applications. The unavailability of such data restricts the advancement of research and development in ABSA for Marathi.

B. Complex Morphological and Syntactic Structures

Marathi is a morphologically rich and syntactically complex language. Words often exhibit different forms based on tense, gender, case, and context, making it challenging to extract aspect terms and sentiments accurately. Additionally, the language's flexible word order and intricate sentence structures further complicate the process of identifying meaningful patterns, requiring sophisticated linguistic and computational methods.

C. Ambiguity in Aspect Terms and Sentiment

Ambiguity in aspect terms and sentiments is a significant challenge in Marathi. A single word or phrase can have multiple meanings depending on the context in which it is used. For instance, the same term might represent different aspects in different domains or could be interpreted differently by individual users. Resolving such ambiguities requires advanced contextual analysis and well-annotated data to train models effectively.

Table 3 . Example of Ambiguity between language

Word	Meaning 1	Meaning 2
□□□□ (Pooja)	A person’s name (commonly used for females)	Offering of prayers to deities
□□□□□ (Vishal)	A person's name (commonly used for males)	Vast or large (used to describe something grand)
□□□(Period)	Time or period (duration)	Color (as black)

D. Limited Availability of Language Processing Tools

The Marathi language lacks robust tools for essential natural language processing tasks, including tokenization, lemmatization, and part-of-speech tagging. These foundational processes are critical for preparing text data for ABSA.Existing tools are often rudimentary or underperform due to insufficient training data, thereby affecting the overall accuracy and efficiency of ABSA systems.

E. Diverse Expressions of Sentiment

Sentiment in Marathi is often conveyed through idiomatic expressions, metaphors, and culturally nuanced phrases, making it difficult to classify sentiments accurately. Furthermore, the language frequently employs implicit or indirect expressions of sentiment, such as sarcasm or subtle negativity, which are harder for models to detect. This diversity in sentiment expression underscores the need for models with advanced contextual understanding to perform well in ABSA tasks for Marathi.

C. (RQ3) WHAT DATASETS ARE AVAILABLE FOR ABSA FOR MARATHI LANGUAGE?

1. AI4 Bharat Marathi Sentiment Dataset

The AI4 Bharat project, which focuses on enabling language technology for Indic languages, provides a Marathi sentiment analysis dataset.Although the dataset primarily focuses on sentiment classification, it can serve as a foundational resource for ABSA tasks. To tailor the database for ABSA, further preprocessing and annotation are required to identify specific aspects and associate sentiments with those aspects. The AI4Bharat dataset is publicly available and can be leveraged for both training and fine-tuning aspect-based models in Marathi.

2. Custom Dataset Creation

Given the scarcity of dedicated ABSA datasets in Marathi, a practical approach is to create custom datasets by annotating Marathi texts. This can be accomplished by collecting reviews, opinions, or social media posts in Marathi related to a specific domain, such as restaurants, products, or services. These texts can then be manually labeled to extract relevant aspects (e.g., "service quality," "taste," "price") and classify the associated sentiment (e.g., positive, negative, or neutral). While this approach is time-intensive, it ensures that the dataset is specifically tailored to the needs of the ABSA task.

3. Indic NLP Corpora

The Indic NLP Corpora, developed as part of the Indic NLP project, provides a variety of text resources for multiple Indian languages, including Marathi. While these resources are not explicitly designed for ABSA, they can be valuable for pretraining language models or for use in text classification tasks, which are integral components of ABSA. The Indic NLP datasets, such as the Indic Sentiment Treebank, contain annotated sentences in Marathi and may be adapted for aspect-based analysis by extending their annotation to include aspect-terms.

4. Social Media Data

Another promising source for Marathi ABSA datasets is social media platforms like Twitter , Facebook and Instagram . By collecting publicly available posts and comments in Marathi, researchers can compile a large, diverse dataset for ABSA. In particular, posts related to specific products, brands, or services can be annotated with aspects (e.g., "delivery," "taste," "customer service") and their associated sentiments. Social media datasets provide a wealth of real-time data, capturing public opinion and consumer sentiment on a wide range of topics. However, challenges such as noise, language variety, and sarcasm in social media content require careful preprocessing and annotation.

5. Crowdsourced Datasets

In the absence of existing Marathi ABSA datasets, crowdsourcing represents a viable option for dataset creation. Platforms such as Amazon Mechanical Turk, or regional Indian crowdsourcing services, can be used to gather annotated texts. The process would involve presenting workers with sentences in Marathi and asking them to identify aspects and classify sentiment for each aspect. Crowdsourcing allows for the rapid collection of data at scale and can help create a diverse dataset that represents a wide variety of opinions and domains.

IV IMPLICATION

A.THEORETICAL IMPLICATIONS

The research into Aspect-Based Sentiment Analysis (ABSA) for Marathi provides valuable insights into advancing natural language processing (NLP) for regional languages. While global languages like English benefit from robust tools and techniques, Marathi and other regional languages still face significant challenges due to limited resources and technological advancements. This study demonstrates the potential of BERT-based models and fine-tuning strategies to address these gaps, offering promising approaches to handle Marathi's unique linguistic features effectively. Key technical breakthroughs include the integration of BERT with task-specific layers, such as SQuAD, which has shown a significant improvement in processing Marathi. The bidirectional nature of BERT has been particularly effective in capturing the contextual nuances of the language. Additionally, multi-task learning approaches incorporating tasks like MNLI, NER, and question-answering capabilities have proven to enhance overall performance, paving the way for more advanced models tailored to regional languages.

One of the study's significant findings is the identification of critical resource gaps in Marathi NLP. These include a lack of annotated datasets, limited availability of Marathi-specific pre-trained models, and an absence of standardized evaluation metrics for ABSA tasks. Addressing these resource limitations is essential for fostering further progress in this domain. The research also emphasizes the challenges posed by Marathi's rich morphology, flexible word order, and context-specific sentiment variations, which demand innovative and specialized approaches beyond conventional ABSA methods. Looking ahead, the study outlines several priority areas for future development. These include creating Marathi-specific pre-trained models, establishing standardized datasets for ABSA, enhancing contextual understanding in sentiment analysis, and integrating domain-specific knowledge to improve accuracy. These advancements would not only strengthen Marathi language processing but also serve as a framework for developing NLP solutions for other regional languages. By bridging the gap between global and regional languages, this research marks a significant step toward more inclusive and effective NLP systems, ultimately fostering greater representation of regional languages in the digital world.

B. PRACTICAL IMPLICATIONS

The practical implications of this research emphasize several essential considerations for advancing Marathi Aspect-Based Sentiment Analysis (ABSA). One of the key challenges identified is the lack of standardized datasets. Although many studies report strong performance metrics such as accuracy, precision, recall, or F1 scores, the use of non-standardized datasets makes it difficult to compare results across different models reliably. This inconsistency highlights the pressing need for benchmark datasets that can serve as a standard for evaluating Marathi ABSA systems. Another important insight is the untapped potential for developing new datasets in the Marathi language. Currently, there is a notable bias toward Indo-Aryan language families and English code-mixed data, leaving pure Marathi datasets underrepresented. Addressing this gap would contribute significantly to creating reliable evaluation metrics for Marathi ABSA systems and improving their overall performance and generalizability. The research also offers a practical framework for developing ABSA models by outlining the essential steps involved in building robust systems. Effective preprocessing of raw data is critical to ensure that relevant information is selected for analysis.

Systematic data annotation plays a key role in providing accurate and consistent labels, which are crucial for training high-quality models. Feature extraction enables the identification of important textual elements, while the classification modeling process requires careful consideration of training scenarios to achieve optimal results. This framework provides a structured and practical guideline for researchers, especially those new to the field, offering a clear pathway for developing reliable and scalable ABSA systems for Marathi. By addressing issues such as dataset standardization and resource development, the study lays the groundwork for future advancements in Marathi language processing, enabling broader and more effective applications of sentiment analysis in regional languages.

V. CONCLUSION

The systematic literature review on Aspect-Based Sentiment Analysis (ABSA) for the Marathi language analyzed 100 research papers using a rigorous selection process. After removing 20 duplicate studies, applying inclusion and exclusion criteria that eliminated 10 papers, and assessing quality standards that excluded an additional 55, a total of 15 high-quality studies were selected for detailed analysis. These comprised 4 journal publications and 11 conference papers, providing valuable insights into the current state of Marathi ABSA research. The findings indicate that BERT-based models, particularly when enhanced with task-specific layers, have shown significant promise for Marathi sentiment analysis. However, the research also highlights key challenges, such as the limited availability of annotated datasets, the scarcity of Marathi-specific pre-trained language models, and the inherent complexity of Marathi's rich morphological structure and syntax. Despite these challenges, the proposed system architecture—combining BERT models with task-specific fine-tuning and integrated layers for various NLP tasks—offers a solid foundation for advancing research in this domain. The relatively small number of selected studies (15 out of 100) points to a substantial gap in high-quality research on Marathi ABSA, with journal publications accounting for only 26.67% of the selected papers. This underscores the need for more rigorous and comprehensive studies in this area. Additionally, the predominance of conference papers (73.33%) suggests that the field is still evolving, with ongoing exploration and emerging opportunities for innovation. Looking ahead, this review emphasizes the importance of addressing critical gaps in Marathi ABSA research. Priorities include the development of standardized datasets, the creation of Marathi-specific pre-trained language models, and the improvement of contextual understanding in sentiment analysis. By focusing on these areas and upholding high-quality research standards, future efforts can significantly enhance ABSA capabilities for Marathi and potentially other regional languages, contributing to broader advancements in natural language processing.

VI. REFERENCES

[1] Deshmukh, Sujata, et al. "Sentiment analysis of Marathi language." Int. J. Res. Publ. Eng. Technol.[IJRPET] 3 (2017): 93-97.

[2]Ansari, Mohammed Arshad, and Sharvari Govilkar. "Sentiment analysis of mixed code for the transliterated hindi and marathi texts." International Journal on Natural Language Computing (IJNLC) Vol 7 (2018).

[3] Divate, Manisha Satish. "Sentiment analysis of Marathi news using LSTM." International journal of Information technology 13.5 (2021): 2069-2074.

[4] Kulkarni, Atharva, et al. "L3cubemahasent: A marathi tweet-based sentiment analysis dataset." arXiv preprint arXiv:2103.11408 (2021).

[5] Shelke, Mahesh B., et al. "Lexical Resource Creation and Evaluation: Sentiment Analysis in Marathi." Smart Trends in Computing and Communications: Proceedings of SmartCom 2022. Singapore: Springer Nature Singapore, 2022. 187-195.

[6] Pingle, Aabha, et al. "L3Cube-MahaSent-MD: A Multi-domain Marathi Sentiment Analysis Dataset and Transformer Models." arXiv preprint arXiv:2306.13888 (2023).

[7] Chaudhari, Prasad, et al. "MahaEmoSen: Towards Emotion-aware Multimodal Marathi Sentiment Analysis." ACM Transactions on Asian and Low-Resource Language Information Processing 22.9 (2023): 1-24.

[8] Gaikwad, Ramnath Mahadeo, Rajashri Ganesh Kanke, and Manasi Ram Baheti. "Review on Sentiment Analysis of Marathi Language of Maharashtra."

[9] Joshi, Aditya, et al. "Towards sub-word level compositions for sentiment analysis of hindi-english code mixed text." Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. 2016.

[10] Banik, Nayan, et al. "Survey on text-based sentiment analysis of bengali language." 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT). IEEE, 2019.

[11] Madan, Anjum, and Udayan Ghose. "Sentiment analysis for twitter data in the Hindi language." 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2021.

[12] Dutta, Rakesh, et al. "Aspect based sentiment analysis using multi-criteria decision-making and deep learning under COVID-19 pandemic in India." CAAI Transactions on Intelligence Technology 8.1 (2023): 219-234.

[13] Shah, Parita, Priya Swaminarayan, and Maitri Patel. "Sentiment analysis on film review in Gujarati language using machine learning." International Journal of Electrical and Computer Engineering 12.1 (2022): 1030

[14] Naukarkar, Renuka Ashokrao, and A. N. Thakare. "A Review on Recognition of Sentiment Analysis of Marathi Tweets using Machine Learning Concept." Int J Scientific Research Sci, Eng Tech (IJSRSET) 8.2 (2021): 190-193.

[15] Mhaske, N. T., and A. S. Patil. "Sentence Annotation for Aspect-oriented Sentiment Analysis: A Lexicon based Approach with Marathi Movie Reviews." *Journal of The Institution of Engineers (India): Series B* (2024): 1-9.

[16] Yulianti, Evi, and Nuzulul Khairu Nissa. "ABSA of Indonesian customer reviews using IndoBERT: single-sentence and sentence-pair classification approaches." *Bulletin of Electrical Engineering and Informatics* 13.5

(2024): 3579-3589.

[17] Chakravarthi, Bharathi Raja, et al. "Dravidian Multimodality: A dataset for multi-modal sentiment analysis in tamil and malayalam." *arXiv preprint arXiv:2106.04853* (2021).

[18] Ramanathan, Vallikannu, T. Meyyappan, and S. M. Thamarai. "Sentiment analysis: an approach for analysing tamil movie reviews using Tamil tweets." *Recent Advances in Mathematical Research and Computer Science* 3 (2021): 28-39.

[19] Sharma, Raksha, and Pushpak Bhattacharyya. "A sentiment analyzer for Hindi using Hindi Senti Lexicon." *Proceedings of the 11th International Conference on Natural Language Processing* (2014): 150-155.

[20] Nishiwaki, Yasuto, Masataka Oshima, Kiyota Hashimoto, and Kazuhiko Tsuda. "A Study of Sentiment Analysis based on Specific 6-emotion Category for Thai Language." *Information Engineering Express* 9, no. 2 (2023).

[21] Jonnalagadda, Pravarsha, Krishna Pratheek Hari, Sandeep Batha, and Hashresh Boyina. "A rule-based sentiment analysis in Telugu." *International Journal of Advance Research, Ideas and Innovations in Technology* 2, no. 5 (2019): 387-390.

[22] Maqsood, Zoya. "Weakly supervised learning for aspect-based sentiment analysis of Urdu Tweets." In *Proceedings of the 8th Student Research Workshop associated with the International Conference Recent Advances in Natural Language Processing*, 78-86. 2023.

[23] Akhtar, Md Shad, Asif Ekbal, and Pushpak Bhattacharyya. "Aspect based sentiment analysis: category detection and sentiment classification for Hindi." In *Computational Linguistics and Intelligent Text Processing: 17th International Conference, CICLing 2016, Konya, Turkey, April 3–9, 2016, Revised Selected Papers, Part II*, 246-257. Springer International Publishing, 2018.