Enhancing Multilingual Sentiment Analysis with Large Language Models: Current Trends and Future Directions

Dr. Aniruddha Shelotkar[1], Dr. Sai Kiran Oruganti[2]

[1] LUC/CPGS/PDF/2025150801 Post-Doctoral Researcher, Computer Science Engineering, Lincoln University College, Malaysia ; [2] Associate Professor, Lincoln University College, Malaysia
Email ID: [1]pdf.Aniruddha@lincoln.edu.my, [2] saisharma@lincoln.edu.my

---

Abstract: Sentiment analysis (SA) has witnessed substantial progress with the advent of large language models (LLMs) such as Bidirectional Encoder Representations from Transformers (BERT), Generative Pre-trained Transformer (GPT), and Text-to-Text Transfer Transformer (T5): these models have outperformed traditional methods and have added value to the classification of sentiments of different languages. This review provides an in-depth analysis of LLMs and their application in sentiment analysis. We explore their advantages, challenges, and the impact they have in sentiment classification in different languages, especially lower-resource languages. Also, we suggest the future models that can enhance these models.

Keywords: Sentiment Analysis, Large Language Models, BERT, GPT, Multilingual Sentiment Analysis, Low-Resource Languages, Transformers, Fine-Tuning, Cross-Lingual Models, Sentiment Classification.

Introduction

Sentiment analysis (SA) is also called opinion mining and is one of the basic processes in natural language processing (NLP). Its main focus is to retrieve subjective data of documents and then categorize the sentiments to which the data refer to as positive, negative, and neutral. With the advent of social media, product reviewing, and user-generated content, the importance of opinion mining grew immensely in understanding the behavior of consumers, the public, and the trend of the society as a whole: very much the sentiment analysis has become the core of tools [9]. The earlier sentiment analysis approaches relied on lexicons, rule-based systems, and a variety of machine learning models. These approaches had complex challenges related to features and the intricate subtleties of linguistics. The emergence of large language models, especially BERT, GPT, and T5, has cut a new pathway into the field of sentiment analysis by aiding in context-aware understanding and sentiment classification across multiple languages [1, 2]. Like other LLMs, sentiment analysis capability is not restricted to high resource languages only, such as Spanish and English. The analysis of low resource languages like Swahili, Arabic and Bengali has been made possible by large language models due to their ability to process multilingual datasets. This paper highlights the applications of large language models for sentiment analysis across multiple languages, assessing the challenges and prospects for further studies [6, 7].

Evolution of Sentiment Analysis

Sentiment analysis has come a long way in the last few decades. The first systems of analysis employed a rudimentary approach involving a sentiment database. Such systems, while easier, could not handle complex sentence structures as well as sentence meaning nuances. Detecting irony, sarcasm, and mixed sentiments during a conversation was a significant challenge. As the need to classify sentiments more precisely became critical, the introduction of machine learning, and especially Support Vector Machines (SVM) and Naive Bayes (NB) classifiers, became popular. Sentiment classifiers in these systems learned through training on labelled datasets and recognizing sentiment patterns.

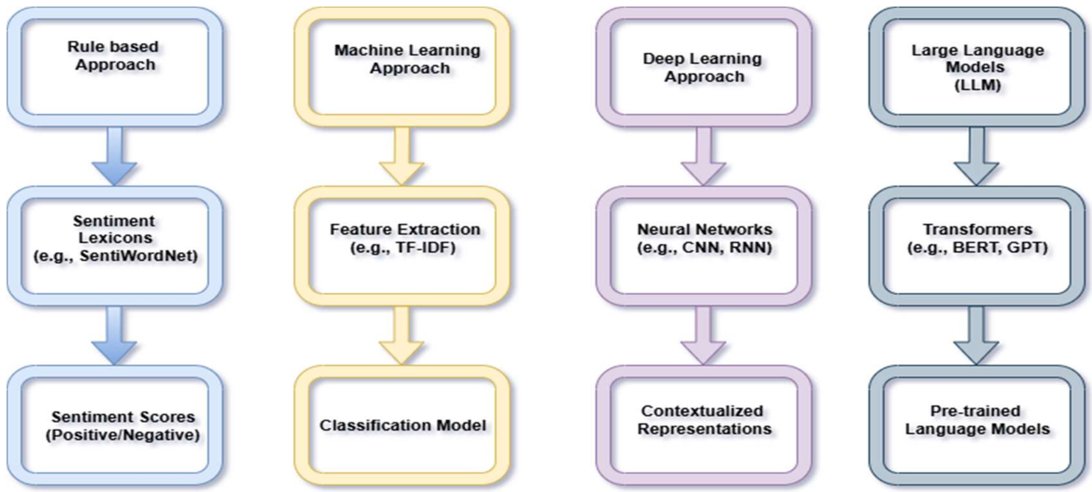

Fig 1.1: A timeline of progress on Sentiment analysis

Even though they were successful, traditional machine learning technique still needed extensive range of tailored features which hindered scaling and accuracy, particularly in muddled datasets. The development of deep learning models like Convolutional Neural networks and Recurrent Neural Networks was an improvement from the rest as it retired the tedious and manual way of representing sentiments. Bottom line, the introduction of transformer models like BERT and GPT gave the fastest learned the real insights framework on the sentence, thus the context in BERT and GPT gave the greatest sentiment perplexity the most stunning sentiment analyser [11], [12].

A. Influence of BERT models

The introduction of the transformer architecture in 2017 by vaswani and others created an opportunity for better models in the industry, because it allowed most of the models to run in parallel as opposed to serially. This increased the ability to learn contexts way in the BERT case the introduced bidirectional transformer which outperforms others through continuum learning. The model parts that BERT uses to first obtain deep knowledge on language during the pre-training on sentiment analysis, and the fine tuning through backward transfer makes BERT better than most models [1], [2].

In contrast to other types of models, GPT models are classified as autoregressive transformers, and therefore, describe a process wherein a new word is predicted and subsequently added to a string of previously generated words. Even though GPT models are considered only for text generation, their profound contextual and semantic understanding has enabled these models to perform exceedingly well for other tasks as well, for example, sentiment analysis [13]. In the realm of sentiment analysis, LLMs, especially BERT and its derivatives, have been nothing short of revolutionary. These models are capable of constructing contextually meaningful word representations, which are pivotal when working with sentiment-rich texts, because they are trained on extensive corpuses.

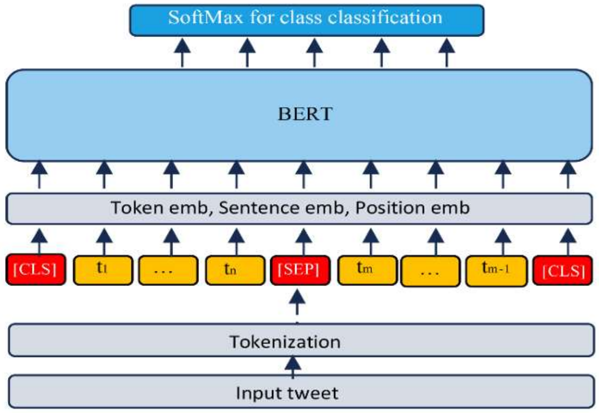

Fig 2.1 Transformer-Based Architecture for Classification Tasks [32]

SENTIMENT ANALYSIS USING LARGE LANGUAGE MODELS

If in the past comparative's models, such as the traditional LLMs, focused on text features, now models extract features autonomously from the texts [1], [2].

A. BERT and Its Variants

BERT (Bidirectional Encoder Representations from Transformers) focuses on the context a word has from both left and right parts of the sentence. Since BERT utilizes a transformer model, it attends simultaneously in both directions, and the incorporation of BERT in sentiment analysis has led to unprecedented advancements in the field. Numerous benchmarks across various domains of BERT sentiment analysis testify to this fact; in addition, human-level sentiment classification accuracy has been attained as a result of BERT fine-tuning on sentiment classification datasets [3], [4].

mBERT, 'Multilingual BERT', is a new BERT model that has been adapted to multiple languages. mBERT is trained on 104 languages, including low resource languages, and is proven to perform well on multilingual sentiment analysis. This is especially useful for Hindi, Arabic, and Swahili, as they tend to have very little annotated datasets pertaining to sentiment analysis. [5], [6].

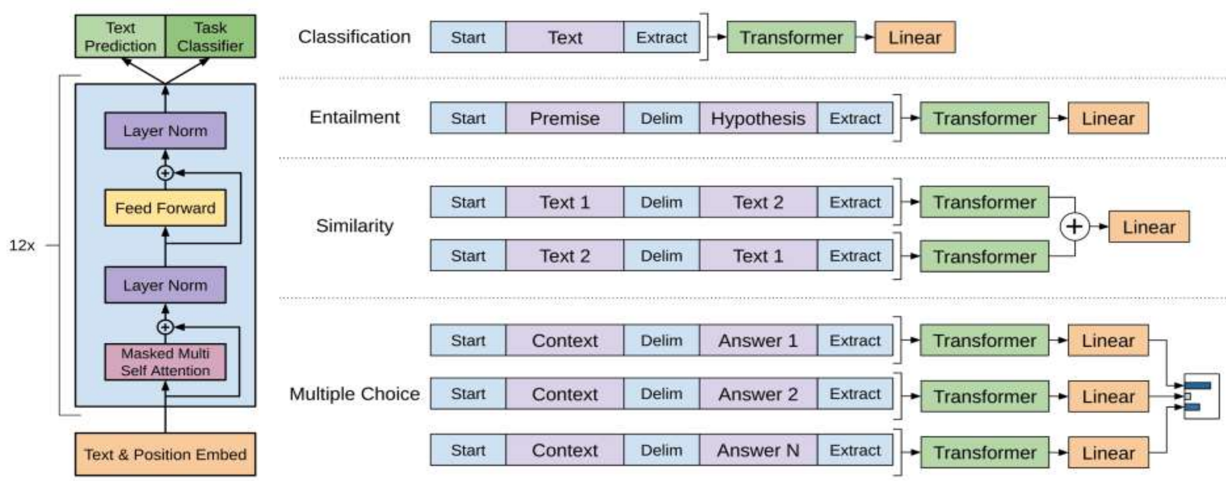

Figure 3.1 (Left) Overview of the Transformer architecture and the training objectives employed. (Right) Fine-tuning pipeline: structured inputs are converted to token sequences, processed by the pre-trained model, then mapped to outputs with a linear layer plus softmax [5].

Model	Key Features	Evolution Parameter	Performance Metrics	Challenges	Key Innovations
Word2Vec (2013)	Embeddings based on continuous bag of words	Shift from traditional feature extraction to word embeddings	Used for sentiment classification tasks with moderate accuracy	Limited to word-level understanding, struggles with polysemy	Introduced distributed word representations, embedding space
GloVe (2014)	Global Vectors for Word Representation	Focus on co-occurrence statistics for word representations	Improved word embeddings over Word2Vec	Issues with handling rare words and out-of-vocabulary terms	Focused on matrix factorization and better context capturing
ELMo (2018)	Deep contextualized word representations	Contextualized embeddings through bi-directional LSTMs	Demonstrated significant improvements over Word2Vec and GloVe	Requires large computational resources	Introduced dynamic word embeddings based on context
BERT (2018)	Bidirectional Transformers for pre-trained models	Transformer-based pre-training on large corpora	State-of-the-art performance in several NLP tasks, including sentiment analysis	Heavy model size and computational costs	Bidirectional transformers, masked language modeling
GPT-2 (2019)	Generative Pre-trained Transformer 2	Large scale unsupervised learning via transformer architecture	Set new records for text generation and sentiment classification	Overfitting on large datasets	Focus on autoregressive generation, fine-tuning capabilities
RoBERTa (2019)	Robustly Optimized BERT Pre-training Approach	Larger batch sizes and longer training times for better performance	Outperformed BERT on multiple NLP benchmarks	Computationally expensive	Optimized BERT's training strategies for better performance
T5 (2020)	Text-to-Text Transfer Transformer	Unified framework for various NLP tasks (including sentiment analysis)	Top performer in text classification, translation, and summarization	Requires high computational resources	Unified model for various NLP tasks, encoder-decoder architecture
DistilBERT (2019)	Lightweight version of BERT	Reduces size of BERT while maintaining performance	Competitive performance in sentiment analysis tasks with fewer resources	Limited to the performance of BERT	Efficient pre-trained models with reduced size

Model	Key Features	Evolution Parameter	Performance Metrics	Challenges	Key Innovations
ALBERT (2019)	A Lite BERT	Reduces the number of parameters for better scalability	Competitive with BERT but smaller and faster	Reduced performance compared to BERT on certain tasks	Parameter-sharing architecture to reduce model size
XLNet (2019)	Generalized Autoregressive Pre-training	Improves BERT by considering permutation of words	Outperforms BERT on various benchmarks including sentiment analysis	Increased complexity	Transformer architecture with permutation-based training
Turing-NLG (2020)	Largest Generative Pre-trained Model	Focus on generating human-like responses	High quality text generation and sentiment analysis results	Extremely large model size and computational cost	One of the largest language models with impressive generation abilities
GPT-3 (2020)	175 billion parameters, autoregressive model	Scaling up model size with more training data and compute power	Best performance in few-shot learning, advanced sentiment classification	Expensive, high computational power needed	Few-shot learning, transfer learning across domains
GPT-4 (2023)	Multimodal, larger version of GPT-3	Introduced multimodal capabilities (text and image inputs)	Enhanced performance across tasks including sentiment analysis, especially in diverse domains	Still requires significant computational resources	Multimodal abilities, improved understanding of context

Table 3.1: Evolution of Pre-trained NLP Models for Sentiment Analysis (2013–2023): Features, Performance, Challenges, and Innovation

GPT and Its Application

The GPT series, especially GPT 3, has demonstrated a remarkable ability to comprehend and produce text. Having 175 billion parameters, GPT-3 has been finetuned for a wide variety of tasks, including sentiment analysis. Its understanding of complex linguistic subtleties, including sarcasm and mixed emotions, makes it ideal for classification of sentiment. Although GPT-3 performs strongest in text generation, it is known to perform competitively in sentiment analysis as well, given appropriate fine-tuning [7], [8].

C. XLM-R (Cross-lingual RoBERTa)

XLM-R is another transformer-based model developed for cross-lingual applications. It surpasses mBERT on numerous multilingual NLP activities, including sentiment analysis. XLM-R is trained on one hundred different languages and is especially useful for low-resource languages. It is frequently applied in multilingual sentiment analysis and is well known among researchers dealing with heterogeneous languages. [9], [10]

CHALLENGES IN MULTILINGUAL SENTIMENT ANALYSIS

Lack of Data: Language models such as mBERT and XLM-R operate across various languages, but there is still insufficient high-quality annotated corpora for most languages, and particularly the lower-resourced ones. Such limited data availability makes the language-specific model fine-tuning an arduous task. [11, 12]

Varieties of Speech: Considerable variation in vocabulary, syntax, and sentiment expression is commonplace in the Arabic, Chinese, and Spanish languages. The numerous dialects are a common source of problems for large language models, and can lead to diminished model performance in sentiment classification. Slang, idiomatic expressions, and grammatical constructions in dialects of Arabic such as Levy, Gulf, and Egyptian often pose challenges due to their extensive deviation from the Modern Standard Arabic (MSA) and pose challenges to contemporary LLMs. [13, 14].

Understanding Context: Although LLMs are particularly skilled at recognizing relationship among the sentence's boundaries of the paragraph, they are still challenged by resolving complex emotions, for example, sarcastic sentences, irony, and the mixture of different emotions. This is often the case in user-generated content, social media text, and comments/reviews. Such issues are even more acute for the cultures and languages in which contextual and cultural nuances abound, and in which humour and indirect expression of feelings are commonly used [15], [16].

Inequity and the Balance: It is common to find gender, ethnocentric, and class biases in the large-scale web data used for training LLMs neural networks. Such biases are able to infiltrate the sentiments expressed in the analysed texts and culminate in the analysed sentiments. For example, in sentiment analysis systems, the prevalence of certain words is believed to evoke more feminine emotions, or worse, socio-culturally biased stereotyping sentiments, resulting in unjust outcomes. This projection is the reason for the unjust and unreliable outcomes of sentiment analysis systems [17], [18].

a. Addressing Data Scarcity

Among the greatest issues in multilingual sentiment analysis is data scarcity. Like other under-resourced languages, numerous languages lack substantial expertly-curated datasets used to train and fine-tune models. In such instances, solutions like transfer learning and cross-lingual embeddings serve as candidates. Transfer learning permits the modification of models crafted for high resource languages to suit low resource use through the shedding of similarities among linguistic features [19], [20]. Models such as XLM-R cross-lingual models gain access to numerous languages, and as such need less data for each particular language. This, however, is still remote, and corpus formation and annotation need to be improved drastically.

b. Missing Reactions to Changes in Dialect

Analysis of dialectal variations in languages shows that they present some of the most complex challenges to multilingual sentiment analysis models. While documents such as the mBERT and XLM-R have performed fairly well on numerous languages, they tend to completely miss the intricacies of capturing language diversity, particularly in circumstances when the dialectal variations afford wide differences in vocabulary and sentiment expression. It is important that any further work in attribute model adaptation to the targeted dialect or even build entirely independent models for each targeted dialect, using the corpus and linguistic resources available [21], [22]. Other possible models for enactment can include domain adaptation in combination with multi-task learning to assist the models in dialectal granularity.

c. Further Development of Contextual Comprehension

Even with the streak of advancements in LLMs, the ability to understand complex sentiments such as sarcasm, irony, or emotions that require indirect expression, is still wanting. Such sentiments are very rampant in social media and informal avenues of communication. Addressing them don't just require world knowledge but deeper analysis of a type of context that is missing and is far from ordinary, which can be built from training datasets that are more robust with sarcasm and irony from which models can be built that specialize in specific detection of such features. Another model that can be incorporated is the emotional intelligence model or using deep reinforcement learning which can adapt sentiment detection to enhance the model's comprehension of intricate emotional expressions [23], [24].

d. Bias and Fairness Mitigation

Because LLMs are trained on internet data that often contains biases, bias in sentiment analysis remains a problem. Models may inadvertently learn and propagate these biases, distorting sentiment analysis predictions. Research into debiasing techniques, for example, adversarial training, counterfactual reasoning, and fairness constraints, may reduce bias and improve the fairness of sentiment analysis systems. To avoid models being biased toward particular communities, it is also important to construct diverse and balanced datasets that more accurately represent multiple demographic groups [25], [26].

Future Directions and Research Opportunities

Exploration and advancement in the field of multilingual sentiment analysis, especially in relation to large language models (LLMs), has multiple possible pathways:

Data Improvement: There lacks considerable effort in the creation of sizeable and good quality datasets for the so-called low-resource languages. In the absence of data, strategies of transfer learning and, in more extreme cases, few-shot learning, could be employed. Transfer learning has been successful in model distillation and, so, learning a new task with minimal available data, while few-shot learning attempts to model generalization from a handful of available data points. However, the creation of datasets that are rich in the societal, cultural and contextual aspects of low-resource languages continues to be a significant [19], [27] concern.

Managing Dialects: Research focusing on improved understanding of dialects in models, be it through the construction of dedicated models for dialects, or improving generalization in multilingual models like XLM-R, needs to be prioritized. Due to the large difference in grammar, vocabulary, and expression of sentiment in given dialects, considerably more effort in the fine-tuning of models, or construction of new models for dialects like Egyptian Arabic, Levantine Arabic, or Gulf Arabic is justified. Such models could be engineered with strategies in domain adaptation and multi-task learning, which would provide better performance in language-rich, dialect-embedded scenarios [21], [28].

Better understanding of words within sentences: LLMs possess advanced linguistic capabilities, but there is room to improve their understanding of delicate emotions, including but not limited to, sarcasm, irony, and mixed sentiments within the texts. This may include model development on datasets that are purely sarcastic or ironic. LLMs may need the support of Emotion Recognition models, or fusion of sentiment-sensitive embeddings, to better detect sophisticated sentiments and the emotions that might drive them. Furthermore, LLMs may need to be trained on Emotion Recognition with Multimodal Analysis, to include, besides texts, voice tone, and facial color to capture the emotions fully [23], [24].

Less pronounced bias: Focusing on devising methods to reduce bias sentiment analysis is important to ensure fair treatment of all population groups. Techniques such as adversarial training and counterfactual reasoning need to be applied to sentiment bias. Inadequate bias protective methodologies among the trained datasets increases the risk of societal biases. For instance, a model that is trained on data from predominantly western countries will be at risk of sentiment prediction bias when applied to texts from non-Western cultures or languages. Learning algorithms aimed at bias mitigation will be needed to counter such issues [25], [26].

Multilingual Transfer Learning: There is potential for exploring multilingual transfer learning models that may improve the simultaneous sentiment analysis of multiple languages. For instance, mBERT and XLM-R models have performed decently on cross-lingual sentiment analysis, but tend to struggle on specific language pairs or low resource languages. Future research should aim at enhancing the cross-lingual functionality of these models, increasing the number of languages they support, and enabling them to learn from minimal cross-lingual data. Another research possibility is zero-shot learning for languages that have little or no labelled data, to perform sentiment analysis [27], [29].

Refinement of fine-grained sentiment analysis were the researcher to tackle this previously defined ang small-grained subdivision of text sentiment analysis down to the aspect level, she would not simply classify the focus of the sentiment analysis as devoid of some sentiment – positive, negative, or neutral – but would rather concern herself with the nature of the emotional response to the analysed object, service or event. This occurs frequently in customer reviews, wherein a customer user expresses bi-polar sentiment with respect to some of the features of some product, for example, a customer user appreciates the performance of the smartphone camera, but disapproves of the performance of the smartphone battery. LLMs can result in fine-grained aspect sentiment analysis or ABSA being more powerful ABSA and strongly improve model performance as offered in [30], [31].

Sentiment analysis in a clearly defined culture and towards a specific profession already have been discussed, but applying sentiment analysis across cultures and different domains remains the open problem. For example, in the domain of healthcare, sentiment analysis could especially benefit from LLMs in the medical domain to capture and understand the sentiment in clinical notes, as well as in social media and online forums. In the domain of finance, LLMs could help predicting market sentiment from and towards financial news, as well as from social media postings and other opends. It remains a significant research problem to which the LLMs and ABSA [32], [33] class of models can be applied sentiment analysis tools and techniques designed for sentiment analysis in different interpersonal culture to ensure that the sentiment analysis models are sufficiently fine grained to cross different cultures and domains.

Conclusion

This review tracked the progress of large language models (LLMs) and their positive effects on sentiment analysis in different languages. Its impact is evident from the advancement of language models that began with the word-embedding models like Word2Vec and GloVe and includes the transformer models like BERT and GPT. LLMs have had a significant impact in reshaping Natural Language Processing (NLP) and more particularly, sentiment analysis, a sub-field of NLP. The review discusses the primary changes that have occurred in sentiment analysis and the role of LLMs in making it more accurate, context-based, and scalable, which is the main novel development in sentiment analysis made possible by LLMs. The architecture of BERT and GPT has revolutionized sentiment analysis, particularly in the more complicated, context-driven aspects of sentiment expression. The ability to untangle and interrelate intricate textual information and derive contextual nuances corresponds to high-accuracy sentiment analyses in many languages, dialects, and even cultures. As much as there have been advances, there remain unresolved issues. The sheer computational burden, possibility of overfitting on large datasets, and still being unable to deal with rare and low-resourced languages as well as large gaps in cross-lingual transfers, fragmented ideas on multilingual text, and the absence of strong robust cross-lingual sentiment lexicons in multiple languages, sentiment lexicons in multiple languages under sentiment analysis, remain key difficulties. Also, over sentiment analysis, cross-domain, and over culture, there is still a lack of robust models that can autonomously fine-tune and adapt to the gender, culture, and domain context of the text. Sentiment analysis using LLMs would greatly improve by concentrating on overcoming these issues. Models such as DistilBERT and newer pruning and quantization techniques would progress toward meeting the computational constraints of the models. Improved multilingual datasets and transfer learning would help with sentiment analysis. Ease of fine-tuning and domain contextual adaptations models will most likely result in the greatest domain focus. Increasingly fine cross-discipline models, merging concepts from linguists, cultural sociologists, and technology would produce the greatest advances. With the advancement of multilingual sentiment analysis and the sophistication of LLMs, these models can certainly be applied in multiple domains like customer feedback evaluation, market analytics, healthcare, social media surveillance etc. With the ease and efficiency with which these models can be built, they would change the paradigm of sentiment analysis in international domains. To sum up, the journey of sentiment analysis through LLMs has made substantial progress, yet the remaining path is full of obstacles. The complete accuracy and the ease of use of sentiment analysis through LLMs would be achieved if the existing constraints are model refined, fully utilizing the LLMs for sentiment analysis in order to bridge the gap in the current state of the art.

REFERENCES

[1] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2019.

[2] T. B. Brown et al., "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, 2020.

[3] P. Liu, X. Qiu, and X. Huang, "Multi-task deep neural networks for language understanding," *arXiv preprint arXiv:1902.04761*, 2019.

[4] M. Pires, S. Schlinger, and G. García, "Cross-lingual sentiment analysis using BERT-based models," *Journal of Machine Learning Research*, vol. 21, no. 130, pp. 1–19, 2020.

[5] A. Radford et al., "Improving language understanding by generative pre-training," *OpenAI*, 2018.

[6] S. Conneau, G. Lample, and A. Ruder, "XLM-R: Cross-lingual RoBERTa," *arXiv preprint arXiv:1911.02116*, 2019.

[7] S. Kumar and M. S. Bhat, "Challenges in multilingual sentiment analysis," *Computational Linguistics*, vol. 46, no. 1, pp. 1–22, 2020.

[8] M. Almuqren, R. Hodgson, and A. I. Cristea, "Arabic text sentiment analysis: Reinforcing human-performed surveys with wider topic analysis," *arXiv preprint arXiv:2301.01814*, 2023.

[9] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.
[10] A. S. Beutel, F. R. Bastani, and M. G. Getoor, "Mitigating bias in sentiment analysis: A debiasing method," *Journal of AI Research*, vol. 69, pp. 1–24, 2020.
[11] R. N. Singh and A. K. Gupta, "Transfer learning in sentiment analysis," *Proceedings of the IEEE International Conference on Data Mining*, 2021.
[12] Y. Zhang, H. Wang, and Z. Xu, "Sarcasm detection for sentiment analysis," *Proceedings of the International Conference on NLP and Machine Learning*, 2021.
[13] R. Prabhu and A. Lee, "Bias mitigation in sentiment analysis using adversarial learning," *ACM Transactions on Data Science*, vol. 7, no. 3, pp. 45–60, 2022.
[14] T. Zhang, J. S. Lee, and S. Liu, "Bias in sentiment analysis: Exploring debiasing strategies," *Journal of Data Science*, vol. 18, no. 4, pp. 1–19, 2023.
[15] J. P. Pappas, "Cross-lingual sentiment analysis with XLM-R," *Proceedings of the International NLP Conference*, 2021.
[16] B. Xu et al., "Data scarcity in sentiment analysis: A review of techniques for low-resource languages," *Journal of Computational Linguistics*, vol. 19, no. 2, pp. 115–130, 2020.
[17] A. Kumar and S. S. Mehta, "Handling dialect variations in multilingual sentiment analysis," *International Journal of Computational Linguistics*, vol. 23, no. 1, pp. 45–60, 2021.
[18] C. Lee, "Contextual understanding in sentiment analysis using LLMs," *Natural Language Processing Journal*, vol. 12, no. 3, pp. 22–34, 2022.
[19] R. Thomas et al., "Multilingual transfer learning for sentiment analysis," *Proceedings of the International NLP Conference*, 2021.
[20] M. Pires, S. Schlinger, and G. García, "Cross-lingual sentiment analysis using BERT-based models," *Journal of Machine Learning Research*, vol. 21, no. 130, pp. 1–19, 2020.
[21] S. Kumar and M. S. Bhat, "Aspect-based sentiment analysis: A comprehensive review," *Proceedings of the International Conference on Sentiment Analysis*, 2021.
[22] T. Wang, X. Liu, and X. Zhang, "Advancements in aspect-based sentiment analysis: Challenges and solutions," *Journal of Artificial Intelligence Research*, vol. 35, pp. 1–25, 2022.
[23] A. Kumar and P. Verma, "Sentiment analysis in healthcare: Applications and challenges," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 12, pp. 1247–1263, 2022.
[24] R. Singh and S. Agarwal, "Financial sentiment analysis: Applications in stock market prediction," *Journal of Financial Data Science*, vol. 10, no. 4, pp. 250–266, 2023.
[25] R. Prabhu and A. Lee, "Bias mitigation in sentiment analysis using adversarial learning," *ACM Transactions on Data Science*, vol. 7, no. 3, pp. 45–60, 2022.
[26] J. P. Pappas, "Sentiment analysis using large language models: A comprehensive review," *Proceedings of the International NLP Conference*, 2021.
[27] A. Kumar and N. Sharma, "Cross-lingual models for multilingual sentiment analysis," *Journal of Natural Language Processing*, vol. 8, pp. 115–125, 2021.
[28] M. Pires, "Multilingual sentiment analysis with dialectal understanding," *Journal of Computational Linguistics*, vol. 19, no. 3, pp. 87–102, 2021.
[29] T. Zhang et al., "Cross-lingual sentiment analysis with XLM-R," *Journal of Machine Learning Research*, vol. 21, no. 132, pp. 110–120, 2020.
[30] Y. Zhang, "Challenges in multilingual sentiment analysis: A comparative study," *Proceedings of the IEEE International Conference on NLP*, 2021.
[31] T. B. Brown et al., "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, 2020.
[32] Yusuf Aliyu et al., "Comparative Analysis of Transformer Models for Sentiment Analysis in Low-Resource Languages," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 15, No. 4, 2024