WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Multilingual Sentiment Analysis using Deep Learning Approach

Yuktha Bhushan H[1], Yasha B P[2], Visva N[3], Smit Zaveri [4], Pratibha H [5], Kala KU[6]

[1,2,3,4,5,6] **Jain Deemed to be University, Banglore,560041**

[1]**yukthayathin55@gmail.com**, [2]**yashabegur@gmail.com**, [3]**msgtovisva@gmail.com**, [4]**smitzaveri1003@gmail.com**, [5]**pratibhaasanj@gmail.com**
[6]**kala.ku@jainuniversity.ac.in**

ABSTRACT:

This paper explores sentiment analysis in Natural Language Processing (NLP) with a focus on employing BERT (Bidirectional Encoder Representations from Transformers) models. Sentiment analysis involves identifying the emotional tone in textual content and categorizing it as positive, negative, or neutral, with applications in various domains like social media monitoring and customer feedback analysis employing BERT for multilingual sentiment analysis showcases remarkable potential in capturing nuanced sentiments across diverse languages. Its contextual understanding and pre-trained language representations enable effective sentiment classification, transcending linguistic boundaries. By leveraging BERT, we enhance the accuracy and adaptability of sentiment analysis models, facilitating cross-cultural understanding and market insights.

Keywords: Sentiment Analysis, Natural Language Processing (NLP), BERT (Bidirectional Encoder Representations from Transformers), Text Classification, Multilingual Analysis.

INTRODUCTION

Sentiment analysis, a fundamental component of Natural Language Processing (NLP), aims to decipher the sentiment or emotional inclination conveyed in textual data. In today's digital era, with the proliferation of online communication platforms, sentiment analysis has emerged as a crucial tool for understanding public opinion, customer sentiment, and market trends. The process involves analysing text to discern whether the expressed sentiment is positive, negative, or neutral, enabling organizations to extract valuable insights from large volumes of textual data. The significance of sentiment analysis spans across diverse domains, including business, politics, social sciences and finance. In business settings, understanding customer sentiment is paramount for enhancing products or services, maintaining customer satisfaction, and making informed business decisions. Similarly, sentiment-analysis facilitates brand monitoring, reputation management, market research, and competitive analysis. Moreover, sentiment analysis holds relevance in political and social contexts, aiding policymakers, researchers, and analysts in gauging public opinion, tracking societal trends, and informing decision-making processes.

BERT, an advanced NLP framework based on transformer models, has revolutionized the field of sentiment analysis by offering state-of-the-art performance in understanding contextual language semantics. Leveraging pre- trained BERT models, researchers and practitioners have developed innovative techniques for sentiment analysis, ranging from fine- tuning BERT models to employing ensemble methods. This paper comprehensively explores these techniques, shedding light on their methodologies, applications, and potential implications for sentiment analysis tasks.Through an in-depth analysis of sentiment analysis techniques with 8 different BERT models, this paper aims to provide researchers, practitioners, and policymakers with valuable insights into harnessing the power of BERT for sentiment analysis tasks. By elucidating the intricacies of sentiment analysis using BERT and highlighting itstechniques and methodologies, this paper seeks to contribute to the advancement of sentiment analysis research and its practical applications in diverse domains.

APPLICATIONS OF MULTILINGUAL SENTIMENT ANALYSIS

Multilingual sentiment analysis, leveraging advanced techniques like BERT, has become a crucial tool in our interconnected world, offering insights across language barriers. Its applications span diverse sectors, from global business strategies to international policy-making, providing a nuanced understanding of emotions and opinions in multiple languages simultaneously. This technology bridges linguistic divides, enabling organizations and researchers to comprehend global trends, customer feedback, and public sentiment with unprecedented depth and cultural sensitivity.

Global Brand Monitoring and Customer Support: Businesses may monitor how their brand is perceived and receive feedback from customers in a variety of languages and cultural contexts by using multilingual sentiment analysis. With the help of this application, companies may determine how consumer sentiment varies by region, rank replies according to urgency and sentiment, and adjust their tactics for a range of markets. Through the analysis of multilingual customer support interactions, businesses can enhance their worldwide service quality and responsiveness.

Social media and Market Research: In the domain of social media and market research, multilingual sentiment analysis is instrumental in discovering trends, viewpoints or emerging issues that prevail across language barriers.It supports researchers and marketers gaining insights on how a product is received in different cultures, cultural variations of the consumer preferences or detection of viral content and potential crises. This application is particularly valuable for global companies seeking to understand and engage with diverse consumer bases.

Political and financial sentiment tracking: This is another area where multilingual sentiment analysis has a significant impact. Political and financial sentiments indicate how the public thinks about their leaders or a country's GDP, respectively. It enables analysts to track public opinion on political topics, evaluate worldwide reactions to news events, and forecast stock market movements in a multilingual stream using live sentiment extraction from news and social media. Well-suited for policymakers, investors, and others working on a global scale as a tool for gaining access to public opinion on issues of interest.

Content Recommendation and User Experience: Multilingual sentiment analysis helps in improving personalization and user experience on international platform by assisting content recommendation. This can leadto more tailored recommendations for the user and better engagement with providers of content by aggregating data around sentiment in users' preferred languages. Applications reach across many industries; from streaming services and e-commerce platforms to news aggregators, these providers can now deliver emotionally relevant content for global audiences.

Healthcare and Tourism Feedback Analysis: Multilingual sentiment analysis is increasingly important in healthcare and tourism sectors for analyzing patient and traveler feedback across different systems and destinations. It helps identify cultural differences in healthcare perceptions, improve medical services based on multilingual patient experiences, and enhance tourism offerings by understanding traveler sentiments about destinations in multiple languages. This application contributes to improving global healthcare standards and enhancing international travel experiences.

Cross-lingual Trend Prediction and Academic Research: In the realms of trend prediction and academic research, multilingual sentiment analysis offers powerful insights. It allows researchers and analysts to identify emerging global trends by analyzing sentiment across multiple languages, potentially predicting the global adoption of products or ideas. In academia, it enables the analysis of sentiment in multilingual scholarly discourse,facilitating cross-cultural comparisons of attitudes towards scientific topics and enhancing international collaborative research efforts.

TECHNIQUES USED IN PROJECT

BERT (Bidirectional Encoder Representations from Transformers) forms the core of multilingual sentiment analysis projects. This powerful pre-trained language model captures bidirectional context effectively. In this study, we employed nine different BERT-based models to comprehensively analyze their effectiveness in sentiment analysis tasks. These models include DistilBert, BERT (base), ALBERT, RoBERTa, spanBERT, Multilingual BERT, XLM-Roberta (multilingual), mBERT, XLM (multilingual). Each model brings unique architectural advantages, such as reduced model size, multilingual capabilities, or enhanced contextual understanding, making them well-suited for diverse sentiment analysis scenarios. Fine-tuning these pre-trained models for sentiment analysis involves adding a classification layer on top of BERT and training it on sentiment-labelled data. BERT provides deep, contextualized word embeddings, crucial for capturing nuanced language differences.

Multilingual embeddings: It represent words and sentences from different languages in a common vector space. While BERT inherently provides multilingual embeddings, specialized techniques like MUSE (Multilingual Unsupervised or Supervised Embeddings) can be employed for specific tasks or comparisons.

Data augmentation: This proves crucial for languages with limited sentiment-labelled data. Techniques include back-translation, where sentences are translated to another language and back, creating variations. Contextual augmentation using language models replaces words with synonyms while maintaining sentiment. Cross-lingual data augmentation leverages sentiment-labelled data from high-resource languages to bolster low-resource language datasets.

Fine-tuning strategies: It significantly impact BERT's adaptation to multilingual sentiment tasks. Gradual unfreezing begins by fine-tuning top layers and progressively unfreezes lower modules. Discriminative fine- tuning applies varying learning rates to different BERT layers. Language-specific fine-tuning further enhances performance.

Cross-lingual alignment: It ensures consistent sentiment analysis across languages. Adversarial training, which trains a discriminator to distinguish between language-dependent corpora, helps the model learn language- agnostic representations. Aligning representations using parallel sentiment-labeled data, when available, proves beneficial.

These techniques collectively enable the construction of robust multilingual sentiment analysis systems using BERT and deep learning. They address key challenges in cross-lingual work, including handling low-resource scenarios, maintaining performance parity across languages, and leveraging large pre-trained models. Proper implementation and tuning of these approaches unlock the full potential of multilingual sentiment analysis**.** By leveraging the strengths of these models, we explored techniques such as fine-tuning, ensemble methods, and transfer learning to evaluate their performance across multilingual and domain-specific datasets. The comparative analysis of these models provides valuable insights into their applicability and potential for advancing sentiment analysis tasks [4][6][9]

LITERATURE SURVEY

PAPER	REFERENCE PAPER	AUTHOR(s)	METHODOLOGY	TECHNIQUES USED
Paper 1	Multi class sentiment analysis of Urdu text using multilingual BERT	Lal Khan et.al	Creation of a multi-class Urdu sentiment dataset: 9,312 manually annotated reviews from various domains. Manual annotation into Positive, Negative, and Neutral sentiments. Sentiment classifiers trained using various algorithms and text representations. Models tested: Rule-based, machine learning, deep learning.	Deep Learning Models: CNN-1D, LSTM, Bi-LSTM, GRU, Bi-GRU. Fine-Tuned Multilingual BERT (mBERT): BERT embeddings forUrdu sentiment classification. BERT word embeddings.
Paper 2	Sentiment Analysis on Multilingual Code-Mixed Kannada Language	Satyam Dutta et.al	Sentimentanalysisof code-mixedKannada text (YouTube comments). Preprocessing: Removing emojis, URLs, numbers, special characters,and lowercase-conversion. Various models applied: machine learning, and transformer-based.	BERT with TensorFlow. BERT with ktrain (wrapper). CNN + Bi-LSTM combination.
Paper 3	Sentiment Analysis for Bengali Using Transformer Based	Anirban Bhowmick et.al	Sentiment analysis on Prothom Alo, YouTube-B, and	Pre-trained Transformers: Multilingual BERT,

	Models		Book-B datasets with two-class and three-class classifications. Compared LSTM, CNN, and GRU layers on transformer models with baselines.	XLM-RoBERTa (XLM-R). Fine-tuning BERT and XLM-R for specific datasets.
Paper 4	An Accurate Approach for Sentiment Analysis on Hindi-English Tweets Based on Bert and pseudo-label Strategy	Wei Bao et.al	Pseudo-labeling: Semi-supervised learning to augment labeled data by predicting labels for test data. Multi-sample Dropout to improve model generalization and avoid overfitting. Cross-validation with five folds.	BERT with Pseudo-labeling. TF-IDF with SGDClassifier

Paper 5	Multi-source multi-domain Sentiment Analysis with BERT-based Models	Gabriel Roccabruna et.al	Multi-source multi-domain approach applied over eight Italian sentiment analysis corpora. Predicts sentiment labels (positive, negative, neutral) across multiple domains. Modified AlBERTo architecture.	BERT-based Fine-tuning. Hyperparameter Optimization using Auto-ML tool Optuna (learning rate, weight decay, warmup ratio, epochs).
Paper 6	Sentiment Analysis on Multilingual code-mixing Text	Anita Saroj et.al	Preprocessing: Conversion to lowercase, removal	BERT_BASE fine-tuned model. ReLU and Sigmoid

	Using BERT-BASE: participation of IRLab@IIT(BHU) in Dravidian-CodeMix and HASOC tasks of FIRE2020		of URLs, numbers, white spaces, punctuation, lemmatization. Tasks: Dravidian-CodeMix (English-Tamil, English-Malayalam), HASOC (Hindi, English, German).	activation. Adam optimizer. Dropout layers and recurrent dropout. Softmax output layer.
Paper 7	Multilingual Sentiment Analysis: Overcoming Challenges in Cross-Language Sentiment Detection with NLP	MinnaaAhmad et.al	Cross-lingual transfer learning with multilingual pre-trained models. Data augmentation, domain adaptation, and rule-based sentiment lexicons applied.	Fine-tuning multilingual pre-trained models (mBERT, XLM-R, mT5). Cross- lingual embeddings. Zero-shot/few-shot learning. Machine translation.

Paper 8	Movie Reviews Sentiment Analysis Using BERT	Gibson Nkhata	Fine-tuning BERT for movie reviews, integrating with BiLSTM. Data augmentation techniques (SMOTE, NLPAUG) and a heuristic algorithm for computing overall sentiment polarity.	BERT for language representation. BiLSTM for sequential data. SMOTE for class imbalance. NLPAUG for data augmentation. Heuristic algorithm for polarity computation
Paper 9	A BERT Framework to Sentiment Analysis of Tweets	AbayomiBello et.al	BERT combined with CNN, RNN, and BiLSTM for sentiment analysis on six Kaggle datasets. Preprocessing	BERT, CNN, RNN, BiLSTM. Data augmentation and preprocessing techniques for enhanced accuracy.

			techniques include null value elimination and special character removal.	

METHODOLOGY

In this research, we explored the effectiveness of nine BERT-based models for multilingual sentiment analysis. The methodology followed a systematic approach comprising data preprocessing, model selection, fine-tuning, evaluation, and comparison across various sentiment analysis scenarios.

1. **Data Preparation:**

Dataset Collection: Sentiment-labelled datasets in multiple languages were sourced to ensure linguistic diversity and enable multilingual sentiment analysis. We are using four different regional languages – English, Kannada, Hindi, Tamil.

Preprocessing: Text data underwent cleaning, including removal of stop words, special characters, and redundant whitespace. Tokenization was performed using model-specific tokenizers to prepare input sequences.

2. **Model selection:**

To evaluate the effectiveness of various BERT-based architectures for multilingual sentiment analysis, we selected nine pre-trained models, each offering unique features and capabilities. These models differ in size, training methodology, and specialization, making them suitable for diverse sentiment analysis scenarios. By leveraging their distinct advantages, we aim to identify the most effective models for multilingual tasks and provide a comprehensive comparison of their performance. Below is a brief overview of each model utilized in this study. Eight BERT-based models were selected for their unique features and capabilities:

DistilBERT: A distilled, lightweight version of BERT designed for faster performance and lower computational costs. It balances efficiency with competitive accuracy, making it suitable for resource- limited settings.

BERT (Base): The foundational model providing bidirectional contextual understanding of text. It serves as a benchmark for various NLP tasks due to its robust and versatile architecture.

ALBERT: A more efficient BERT variant that reduces model size through parameter sharing and factorized embeddings. It is optimized for scalability and memory efficiency in large-scale applications.\

RoBERTa: An improved version of BERT trained with enhanced techniques and larger datasets. It achieves superior performance in tasks requiring rich and nuanced language representations.

SpanBERT: A model tailored for span-based predictions, improving contextual understanding within sentence spans. It is particularly effective for tasks like question answering and coreference resolution.

Multilingual BERT (mBERT): A multilingual model trained on over 100 languages, enabling cross- lingual understanding. It is well-suited for tasks involving diverse linguistic inputs without language- specific adjustments. **XLM-RoBERTa:** A robust multilingual extension of RoBERTa, offering state-of-the-art performance for cross- lingual tasks. It handles multilingual datasets with high accuracy and contextual depth.

XLM: A multilingual model focused on translation and cross-lingual tasks, leveraging parallel datasetsfor training. It is designed for scenarios requiring language alignment and representation.

Multilingual embeddings: Multilingual embeddings play a crucial role in sentiment analysis byenabling models to process and understand text across diverse languages within a shared vector space. Models like Multilingual BERT (mBERT), XLM-RoBERTa, and XLM offer cross-lingual capabilities,allowing for consistent sentiment classification and efficient handling of low-resource languages and code-mixed data. By leveraging these embeddings, our approach fine-tunes pre-trained multilingual models to capture language-specific nuances while ensuring scalability and performance across multilingual datasets. This facilitates effective sentiment analysis in global applications such as social media monitoring, customer feedback analysis, and market research.

3. **Data Augmentation:**

This approach becomes particularly essential for languages with limited sentiment-labelled data, as it enables effective model training even in resource-constrained scenarios. By leveraging pre-trained models and applying fine-tuning techniques, we can overcome the scarcity of labelled data, ensuring robust sentiment analysis across diverse languages and datasets.

Contextual Augmentation: Contextual augmentation is used to replace words with synonyms while keeping the sentiment of the sentence intact, ensuring diversity in training data without losing semantic meaning.

Cross-lingual alignment: It ensures consistent sentiment analysis across languages. Adversarial training, which trains a discriminator to distinguish between language-dependent corpora, helps the model learn language- agnostic

4. Fine Tuning Strategies:

Model Fine-Tuning: A classification head (fully connected layer) was added to each pre-trained modelto enable sentiment classification. Models were fine-tuned on the sentiment-labelled datasets.

Optimizers: Optimizing for cross-entropy loss using AdamW optimizer. Learning rate schedules and early stopping mechanisms were applied to enhance training efficiency and prevent overfitting. Hyperparameter optimization is conducted using tools like Optuna for fine-tuning parameters such as learning rate, weight decay, warmup ratio, and number of epochs. This ensures the optimal performance of the sentiment analysis model.

Language-Specific Fine-Tuning: Fine-tuning strategies are customized for each language, ensuring better adaptation to language-specific characteristics and sentiment expression.

5. Fine Tuning Strategies:

Model Fine-Tuning: A classification head (fully connected layer) was added to each pre-trained model to enable sentiment classification. Models were fine-tuned on the sentiment-labelled datasets.

Optimizers: Optimizing for cross-entropy loss using AdamW optimizer. Learning rate schedules and early stopping mechanisms were applied to enhance training efficiency and prevent overfitting. Hyperparameter optimization is conducted using tools like Optuna for fine-tuning parameters such as learning rate, weight decay, warmup ratio, and number of epochs. This ensures the optimal performance of the sentiment analysis model.

Language-Specific Fine-Tuning: Fine-tuning strategies are customized for each language, ensuring better adaptation to language-specific characteristics and sentiment expression.

6. Evaluation metrics:

These metrics are commonly used in classification tasks to evaluate the performance of a model. Here's what each term represents:

Precision: The ratio of correctly predicted positive observations to the total predicted positive observations. Measures the model's ability to avoid false positives.

Recall (Sensitivity or True Positive Rate): The ratio of correctly predicted positive observations to all actual positive observations. Measures the model's ability to identify all relevant instances (avoid false negatives).

F1 Score: The harmonic-mean of precision and recall, balancing the trade-off between the two. Useful when you need a balance between precision and recall, especially in imbalanced datasets.

Support: The number of actual occurrences of each class in the dataset. Indicates the number of samples for each class, which is important for understanding performance in imbalanced datasets.

Accuracy: The ratio of correctly predicted observations to the total observations. Measures overall correctness of the model.

Macro Average (Macro avg): The average of precision, recall, and F1 scores calculated independently for each class. Treats all classes equally by giving them equal weight, regardless of their size.

Weighted Average (Weighted avg): The average of precision, recall, and F1 scores weighted by the number of true instances for each class (support). Takes class imbalance into account by giving more weight to classes with higher support.

Model evaluation: Hardware resources included GPUs to facilitate efficient model training and inference. These metrics collectively help in understanding the strengths and weaknesses of a classification model, ensuring a comprehensive evaluation. The models were compared based on their accuracy, precision. Recall, F1- score, support, macro average, weighted average, processing speed, and ability to handle multilingual inputs. Attention was given to understanding the trade-offs between model complexity, size, and sentiment analysis performance.

EVALUATION OF SENTIMENT CLASSIFICATION

The performance of sentiment classification models is evaluated using key metrics: Accuracy, Precision, Recall, and F1-Score. These metrics are calculated from the confusion matrix, which shows the relationship between actual and predicted sentiment labels.

The matrix includes True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

Table 1. Confusion Matrix

	Predicted Positives	Predicted Negatives
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Precision = TP/(TP+FP)

Recall = TP/(TP+FN)

F1 = (2×Precision×Recall) / (Precision + Recall)

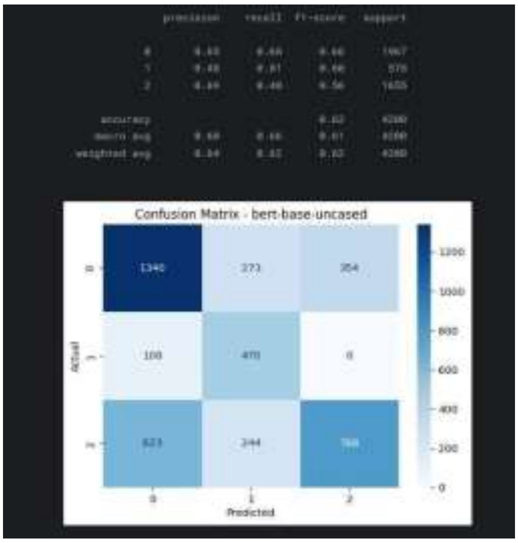
RESULTS

The performance of each model was evaluated using key metrics such as accuracy, precision, recall, F1-score, and support, with the confusion matrix providing an additional layer of insight into the model's classification performance.

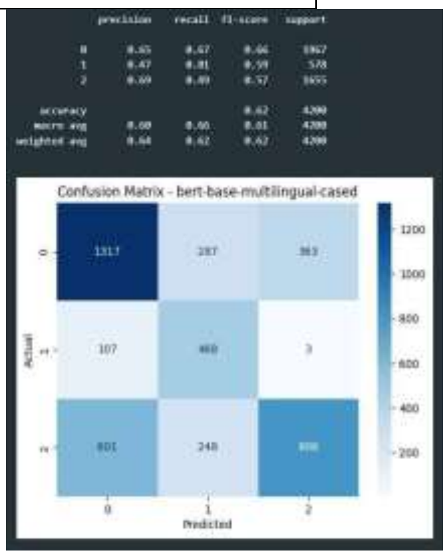
The confusion matrix helped identify the number of true positives, true negatives, false positives, and false negatives, allowing for a detailed assessment of how well each model handled various sentiment categories. Based on these metrics, we summarize the key findings and performance results for each model. The following sections outline the accuracy scores and other performance metrics derived from the evaluation, highlighting the strengths and weaknesses of the models in sentiment analysis tasks.

Table 2. Bert Models and their Accuracy

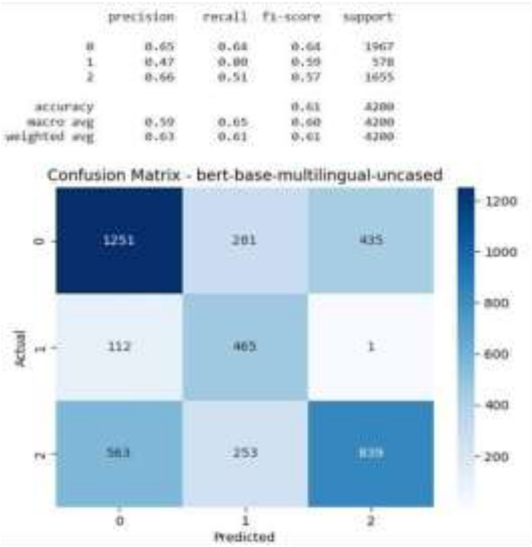
BERT MODELS	ACCURACY
1. Bert-base-uncased model	62%
2. Bert-base-multilingual-cased	62%
3. Bert-base-multilingual uncased	61%
4. Albert-base-V2	66%
5. Spanbert-base-cased	70%
6. Xlm-roberta-base	55%
7. Xlm-mlm-100-1280	47%
8. Distilbert-base-uncased	56%

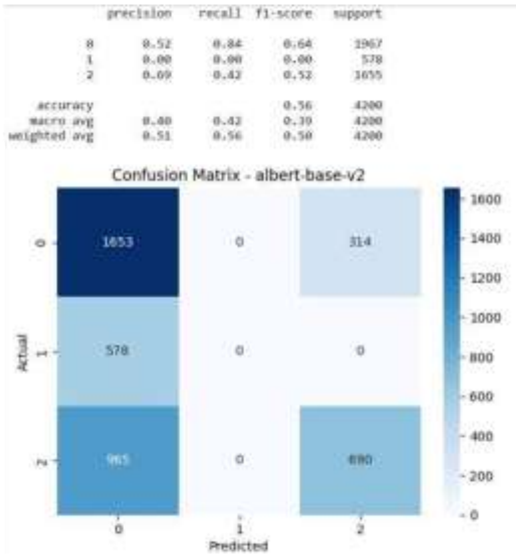


1.Bert-base-uncased model gained 62% accuracy

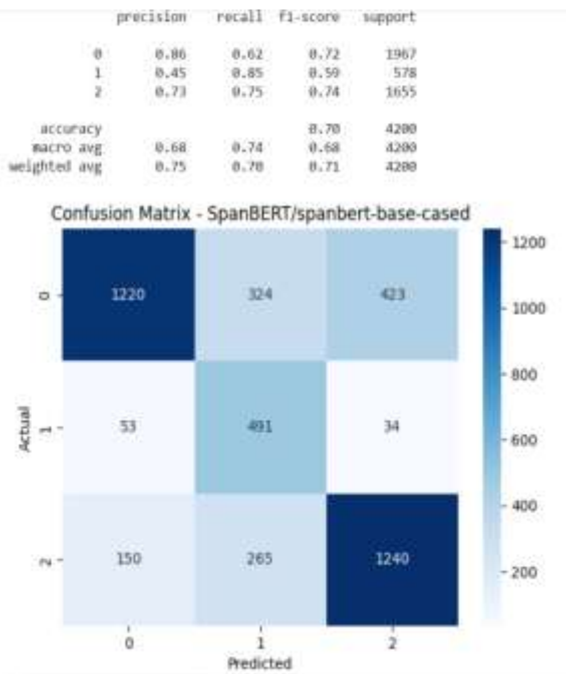


2. bert-base-multilingual cased model gained 62% accuracy



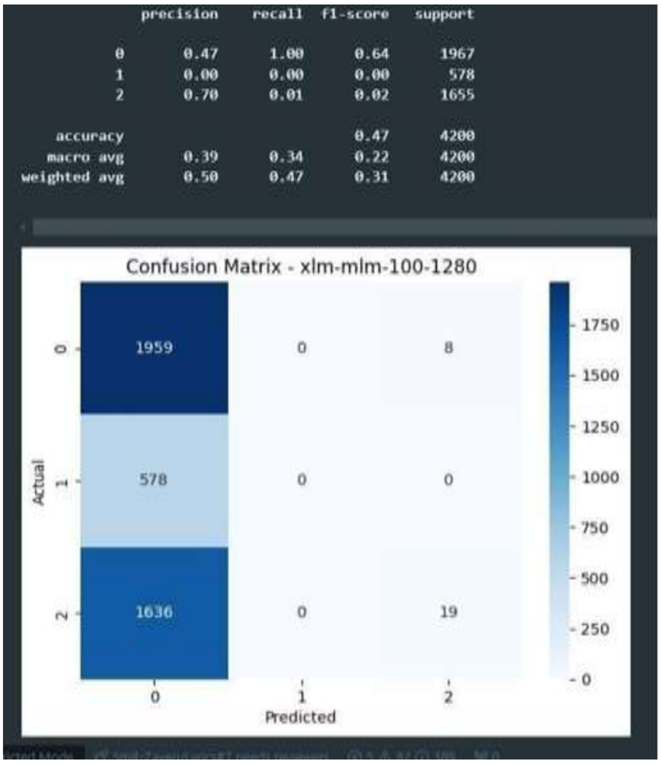


3. bert-base-multilingual-uncased model gained 4. albert-base-v2 model gained 56% accuracy 61% accuracy



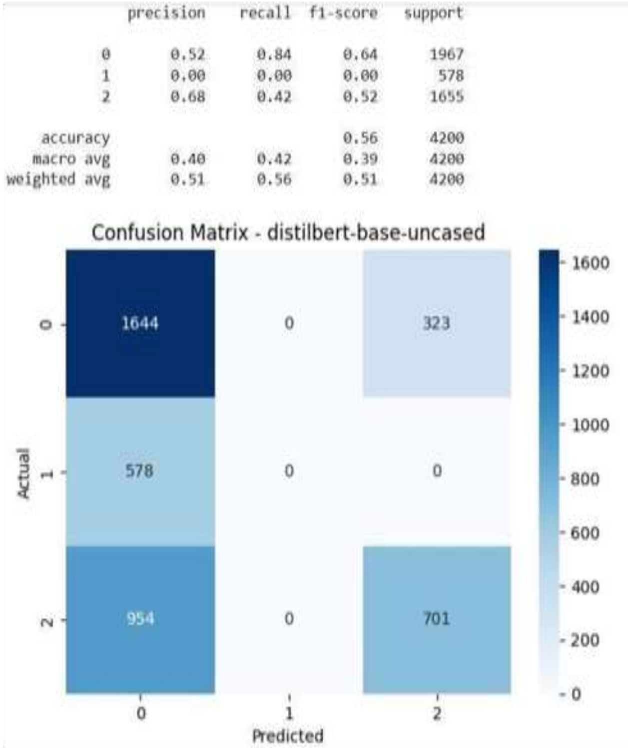
5. xlm-roberta-base model gained 55% accuracy

6. spanbert-base-cased model's accuracy is 70%•



	precision	recall	f1-score	support
0	0.47	1.00	0.64	1967
1	0.00	0.00	0.00	578
2	0.70	0.01	0.02	1655
accuracy			0.47	4200
macro avg	0.39	0.34	0.22	4200
weighted avg	0.50	0.47	0.31	4200

7. xlm-mlm-100-1280 model gained 47% accuracy



	precision	recall	f1-score	support
0	0.52	0.84	0.64	1967
1	0.00	0.00	0.00	578
2	0.68	0.42	0.52	1655
accuracy			0.56	4200
macro avg	0.40	0.42	0.39	4200
weighted avg	0.51	0.56	0.51	4200

8. distilbert-base-uncased model gained 56%

CONCLUSION

In conclusion, this paper provides a survey and comparative study through a systematic exploration of eight BERT variants, including monolingual and multilingual models, the study demonstrates their effectiveness in handling diverse linguistic inputs and extracting meaningful insights from sentiment-labelled data. Research highlights the potential of leveraging BERT-based models for multilingual sentiment analysis. The findings reveal that SpanBERT emerged as the most effective model, achieving the highest accuracy of 70%. Other models, like ALBert and Bert -base, also showed promise in handling multilingual datasets. Evaluation metrics such as accuracy, precision, recall, and F1-score provided a comprehensive assessment of model performance, highlighting the trade-offs between model complexity and efficiency. This research underscores the significance of multilingual embeddings and fine-tuning strategies in enabling sentiment analysis across languages and domains. By addressing key challenges like limited labeled data and language-specific nuances, this study contributes to advancing sentiment analysis research.

FUTURE WORK

Future work could explore integrating these techniques into focusing on optimizing BERT-based models for real-time sentiment analysis on dynamic, large-scale text data from platforms like Twitter and Facebook enhancing the scalability and robustness of sentiment analysis systems for global use case.

REFERENCES

1. Khan, L., Amjad, A., Ashraf, N. et al. Multi-class sentiment analysis of urdu text using multilingual BERT. Sci Rep 12, 5436 (2022). https://doi.org/10.1038/s41598-022-09381-9
2. Dutta, Satyam et al. "Sentiment Analysis on Multilingual Code-Mixed Kannada Language." Fire (2021).
3. Bhowmick, Anirban, Abhik, Jana. "Sentiment Analysis For Bengali Using Transformer Based Models." Proceedings of the 18th International Conference on Natural Language Processing (ICON). NLP Association of India (NLPAI), 2021. 8
4. Bao, Wei, Weilong, Chen, Wei, Bai, Yan, Zhuang, Mingyuan, Cheng, Xiangyu, Ma. "Will_go at SemEval-2020 Task 9: An Accurate Approach for Sentiment Analysis on Hindi-English Tweets Based on Bert and Pesudo Label Strategy." Proceedings of the Fourteenth Workshop on Semantic Evaluation. International Committee for Computational Linguistics, 2020.
5. Roccabruna, Gabriel, Steve, Azzolin, Giuseppe, Riccardi. "Multi-source Multi-domain Sentiment Analysis with BERT-based Models." Proceedings of the Thirteenth Language Resources and Evaluation Conference. European Language Resources Association, 2022.
6. Saroj, Anita, and Sukomal Pal. "Sentiment Analysis on Multilingual Code Mixing Text Using BERT BASE: participation of IRLab@ IIT (BHU) in Dravidian-CodeMix and HASOC tasks of FIRE2020." (2020).

7. Minnaa Ahmad, , Muhammad Shoaib Tahir, Aqsa Shereen, Shehr Bano Zaidi. "MULTILINGUAL SENTIMENT ANALYSIS: OVERCOMING CHALLENGES IN CROSS-LANGUAGE SENTIMENT DETECTION WITH NLP". International Journal of Contemporary Issues in Social Sciences 3. 3(2024): 807–815.
8. Nkhata, Gibson. Movie Reviews Sentiment Analysis Using BERT. MS thesis. University of Arkansas, 2022.
9. Bello, A., S. C. Ng, and M. F. Leung. "A BERT Framework to Sentiment Analysis of Tweets." Sensors.