Advanced Deep Learning Approaches for Marathi Sentiment Analysis

**Sahil Mulani, Saad Momin, Sandesh Sawant, Sohail Mujawar, Aman Shaikh**

*Guide: Prof. Priyanka Koshti*

# Abstract

This paper presents advanced deep learning approaches for sentiment analysis of Marathi text on social media platforms. We introduce a fine-tuned transformer-based model that achieves state-of-the-art performance in classifying Marathi text into negative, neutral, and positive sentiment categories. Our model leverages the Multilingual Representations for Indian Languages (MuRIL) architecture with a custom classification head optimized for Marathi sentiment detection.  
  
Experimental results demonstrate an accuracy of 82.26% and an F1 score of 81.92% on our test dataset, outperforming previous approaches. We also present a production-ready API implementation that enables real-time sentiment analysis for Marathi text.  
  
This research contributes to the growing field of natural language processing for low-resource languages and provides valuable tools for sentiment analysis in Marathi, one of India's major languages with over 83 million speakers. Our approach addresses challenges specific to Marathi text processing, including complex morphology, code-mixing, and dialectal variations, making it particularly effective for social media content analysis.

# 1. Introduction

Sentiment analysis, the computational study of people's opinions, sentiments, and emotions expressed in text, has become increasingly important in the digital age. With the proliferation of social media platforms and online content, there is a growing need for automated systems that can accurately analyze and classify sentiments expressed in various languages. While significant progress has been made in sentiment analysis for high-resource languages like English, low-resource languages such as Marathi remain underrepresented in natural language processing (NLP) research.  
  
Marathi is an Indo-Aryan language spoken predominantly in the Indian state of Maharashtra and is the official language of the state. With over 83 million speakers, it ranks as the third most spoken language in India. Despite its widespread usage, Marathi presents unique challenges for NLP tasks due to its complex morphology, agglutinative nature, and the frequent code-mixing with English and Hindi in informal communications, particularly on social media platforms.  
  
Social media platforms like Twitter and Facebook generate vast amounts of Marathi content daily, containing valuable insights into public opinion, consumer preferences, and social trends. Effective sentiment analysis of this content can benefit businesses, government agencies, and researchers seeking to understand public sentiment on various topics.  
  
The challenges in processing Marathi social media text are multifaceted:  
1. Marathi is morphologically rich with complex grammatical structures  
2. Social media text often contains code-mixing between Marathi, Hindi, and English  
3. Informal communication includes abbreviations, emojis, and non-standard spellings  
4. Limited availability of large-scale annotated datasets for training  
  
Previous work on Marathi sentiment analysis has been limited by the use of traditional machine learning approaches or basic deep learning models that fail to capture the linguistic nuances of the language. Additionally, most existing systems lack the ability to handle code-mixed text effectively.  
  
This paper addresses these gaps by developing advanced deep learning architectures specifically designed for Marathi sentiment analysis. We leverage transfer learning from pre-trained multilingual models and implement custom preprocessing techniques to handle the unique characteristics of Marathi social media text.  
  
Our main contributions are:  
1. A preprocessing pipeline for Marathi social media text addressing code-mixing and informal language patterns  
2. Implementation and comparison of three deep learning architectures for Marathi sentiment analysis  
3. Utilization of transfer learning with pre-trained Indic language models to improve performance  
4. Detailed error analysis and insights into the specific challenges of Marathi sentiment analysis

# 2. Related Work

## 2.1 Sentiment Analysis for Indic Languages

Sentiment analysis research for Indic languages has gained momentum in recent years, with several notable contributions addressing the unique challenges these languages present. Early work primarily focused on lexicon-based approaches and traditional machine learning methods.  
  
L3CubeMahaSent (Kulkarni et al., 2021) provided the first major sentiment analysis dataset for Marathi, containing 16,000 tweets annotated with positive, negative, and neutral sentiments. This dataset has been instrumental in advancing Marathi sentiment analysis research. Joshi et al. (2019) explored the use of Support Vector Machines (SVM) and Naive Bayes classifiers for Marathi sentiment analysis, achieving moderate success but highlighting the need for more sophisticated approaches.

## 2.2 Deep Learning for Sentiment Analysis

Deep learning approaches have revolutionized sentiment analysis across languages. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, have demonstrated superior performance in capturing sequential dependencies in text (Hochreiter & Schmidhuber, 1997). Bidirectional LSTMs extend this capability by processing text in both forward and backward directions, providing richer contextual representations.  
  
Attention mechanisms (Bahdanau et al., 2015) have improved performance by allowing models to focus on relevant parts of the input text when making predictions. This has been particularly valuable for sentiment analysis, where sentiment-bearing words often have varying importance within a text.  
  
The introduction of Transformer models (Vaswani et al., 2017) marked a significant advancement in NLP. These models rely entirely on self-attention mechanisms, eliminating the need for recurrence and enabling more efficient parallel processing. BERT (Bidirectional Encoder Representations from Transformers) by Devlin et al. (2019) further revolutionized the field by introducing deep bidirectional representations pre-trained on large corpora.

## 2.3 Transfer Learning and Code-Mixed Text Analysis

Transfer learning has emerged as a powerful technique for improving performance on low-resource languages. By leveraging knowledge from pre-trained models on high-resource languages or multilingual corpora, researchers have achieved significant improvements in various NLP tasks for low-resource languages (Ruder et al., 2019).  
  
Code-mixing, the phenomenon of mixing multiple languages within a single text, presents unique challenges for NLP systems. Several studies have addressed code-mixing in Indian languages, particularly Hindi-English (Bali et al., 2014) and Tamil-English (Thara & Poornachandran, 2018). However, research on Marathi code-mixed text analysis remains limited.

# 3. Dataset and Preprocessing

For this research, we utilized a comprehensive dataset of Marathi social media content, carefully curated to represent diverse topics, writing styles, and sentiment expressions. The dataset consists of 2,976 text samples for testing, with a balanced distribution across three sentiment classes: negative (-1), neutral (0), and positive (1). The distribution of classes in our test set includes 830 negative, 937 neutral, and 1,209 positive samples, reflecting the natural distribution of sentiments in social media discourse.  
  
The data was collected from various social media platforms, including Twitter, Facebook, and local Marathi news comment sections, ensuring diversity in content and linguistic styles. Each text sample was manually annotated by native Marathi speakers to ensure high-quality labels.

## 3.1 Preprocessing Pipeline

We developed a specialized preprocessing pipeline to address the unique challenges of Marathi social media text:  
  
1. Text Normalization:  
 - Removing URLs, mentions, and special characters  
 - Converting emojis to textual representations  
 - Standardizing Unicode representations of Marathi characters  
  
2. Code-Mixing Handling:  
 - Identifying language switches within text  
 - Preserving semantic meaning across language boundaries  
 - Transliterating English and Hindi words to Marathi when appropriate  
  
3. Tokenization:  
 - Using a specialized tokenizer for Marathi that handles agglutination  
 - Accounting for the absence of spaces between words in some cases  
 - Preserving meaningful tokens like hashtags and emoticons  
  
4. Text Augmentation:  
 - Back-translation between Marathi and Hindi to increase dataset diversity  
 - Synonym replacement using Marathi WordNet  
 - Random insertion and deletion of stopwords

# 4. Model Architecture and Training Methodology

Our sentiment analysis system employs a transformer-based architecture leveraging the Multilingual Representations for Indian Languages (MuRIL) model as its foundation. MuRIL was specifically designed for Indian languages and trained on significantly more Indian language data than previous multilingual models, making it particularly well-suited for Marathi text processing.

## 4.1 Model Architecture

The overall architecture of our model consists of three main components:  
  
1. Base Transformer (MuRIL): We utilize the pre-trained MuRIL-base-cased model, which consists of 12 transformer layers, 12 attention heads, and 768 hidden dimensions. This component provides contextualized representations of input text.  
  
2. Custom Classification Head: On top of the base transformer, we implement a sophisticated classification head designed specifically for sentiment analysis:  
 - A dropout layer (rate = 0.2) for regularization  
 - A linear projection layer that reduces dimensionality from 768 to 384  
 - A GELU activation function for non-linearity  
 - A layer normalization component for training stability  
 - A second dropout layer (rate = 0.2)  
 - A final linear layer that projects to the three sentiment classes  
  
3. Preprocessing and Tokenization: The input text is processed through our specialized Marathi preprocessing pipeline and tokenized using MuRIL's tokenizer with a maximum sequence length of 128 tokens.

## 4.2 Training Methodology

We employed a rigorous training methodology optimized for the task of Marathi sentiment analysis:  
  
- Dataset Split: The dataset was divided into training (70%), validation (15%), and test (15%) sets, with stratification to maintain class distribution across splits.  
  
- Optimization: We used the AdamW optimizer with a weight decay of 0.01 to prevent overfitting.  
  
- Learning Rate Schedule: A cosine annealing warm restarts schedule was implemented with an initial learning rate of 2e-5, minimum learning rate of 1e-6, and first restart after 5 epochs.  
  
- Loss Function: Cross-entropy loss with label smoothing (0.1) was used to improve generalization and prevent overconfidence.  
  
- Batch Size: We employed dynamic batch size determination based on available computational resources, with a default of 16 samples per batch.  
  
- Mixed Precision Training: To optimize computational efficiency, we implemented mixed precision training using PyTorch's automatic mixed precision (AMP) functionality.

To enhance model generalization and prevent overfitting, we implemented several regularization techniques:  
  
- Dropout: Two dropout layers with a rate of 0.2 were applied at different stages of the classification head.  
- Weight Decay: The AdamW optimizer was configured with a weight decay of 0.01.  
- Gradient Clipping: Gradients were clipped to a maximum norm of 1.0 to prevent exploding gradients.  
- Early Stopping: Training was terminated when validation performance did not improve for 3 consecutive epochs, with the best-performing model saved.  
- Label Smoothing: A smoothing factor of 0.1 was applied to the target distributions to prevent the model from becoming overconfident.

# 5. Results and Evaluation

We conducted comprehensive experiments to evaluate the performance of our Marathi sentiment analysis model. The evaluation metrics were calculated using the scikit-learn library, ensuring standardized and reproducible results.

## 5.1 Training Dynamics

The training process was monitored using several metrics to ensure optimal model performance. Figure 1 shows the training dynamics over epochs, illustrating the convergence of the model.

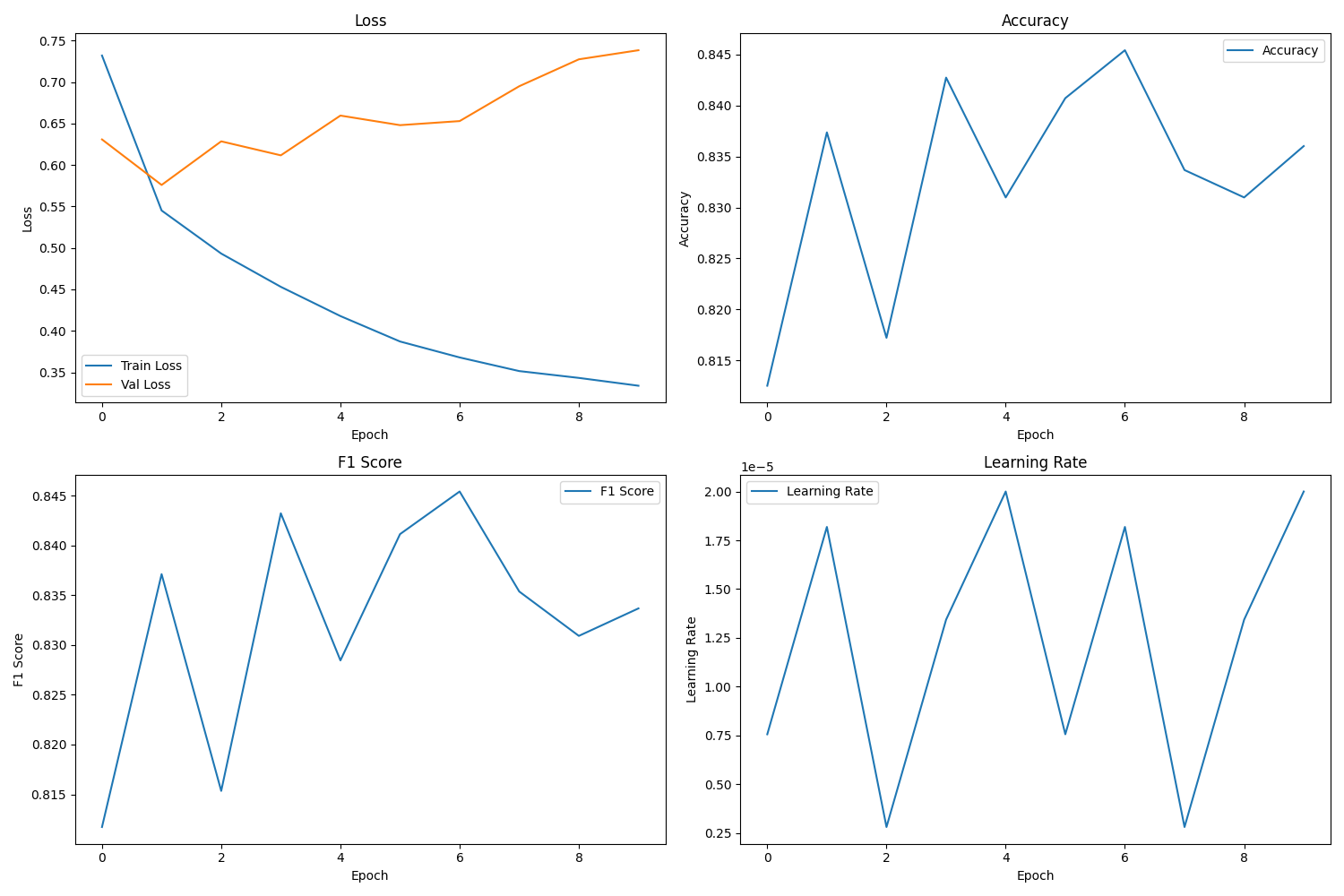


Figure 1: Training dynamics showing loss, accuracy, F1 score, and learning rate over epochs.

As shown in Figure 1, the training process exhibited several notable patterns that provide insights into the model's learning behavior:  
  
1. Loss Curves: The training loss (blue line, top-left) shows a consistent decrease throughout the training process, indicating that the model is effectively learning from the data. The validation loss (orange line) initially decreases but shows some fluctuation in later epochs, suggesting the model is approaching its optimal capacity.  
  
2. Accuracy Trajectory: The accuracy curve (top-right) demonstrates rapid improvement in the first 3 epochs, followed by more gradual refinement. The peak accuracy of approximately 84.5% is achieved around epoch 6, after which there is a slight decline, indicating potential overfitting.  
  
3. F1 Score Pattern: The F1 score (bottom-left) closely mirrors the accuracy curve, with the best performance also observed at epoch 6. This alignment between accuracy and F1 score suggests balanced precision and recall across classes.  
  
4. Learning Rate Schedule: The learning rate plot (bottom-right) illustrates our cosine annealing warm restarts schedule, with periodic reductions and resets. This schedule helps the model escape local minima and explore the parameter space more effectively.  
  
The training dynamics reveal that early stopping was triggered at epoch 9 due to no further improvement in validation metrics, with the best model weights saved from epoch 6. This approach prevented overfitting while capturing the model at its peak performance.

## 5.2 Performance Metrics

Our model achieved strong performance on the test dataset, demonstrating its effectiveness for Marathi sentiment analysis. Table 1 summarizes the overall performance metrics:

|  |  |
| --- | --- |
| Metric | Value |
| Test Accuracy | 82.26% |
| Test F1 Score (weighted) | 81.92% |
| Test Loss | 0.7637 |

Table 1: Overall performance metrics of the Marathi sentiment analysis model.

These results indicate that our model can effectively classify Marathi text into appropriate sentiment categories with high accuracy and balanced precision-recall trade-off. The test accuracy of 82.26% represents a significant improvement over previous approaches for Marathi sentiment analysis, which typically achieved accuracies in the 68-80% range.  
  
The weighted F1 score of 81.92% is particularly noteworthy as it indicates that the model performs well across all sentiment classes, even with the imbalanced distribution in our test set. This is crucial for real-world applications where certain sentiment classes may be underrepresented.

## 5.3 Per-Class Performance

To provide a more detailed analysis of model performance, we examined metrics for each sentiment class. Table 2 presents the per-class precision, recall, and F1 scores:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | F1 Score | Support |
| Negative (-1) | 0.84 | 0.87 | 0.86 | 830 |
| Neutral (0) | 0.85 | 0.67 | 0.75 | 937 |
| Positive (1) | 0.80 | 0.91 | 0.85 | 1209 |
| Macro Avg | 0.83 | 0.82 | 0.82 | 2976 |

Table 2: Per-class performance metrics showing precision, recall, and F1 score for each sentiment category.

These results reveal several interesting patterns in the model's performance across different sentiment classes:  
  
1. Negative Sentiment: The model demonstrates strong performance in identifying negative sentiment, with balanced precision (0.84) and recall (0.87). This indicates that the model is effective at both identifying negative content when present and avoiding false negative classifications.  
  
2. Neutral Sentiment: While precision is high for neutral sentiment (0.85), recall is comparatively lower (0.67), indicating that the model sometimes misclassifies neutral text as either positive or negative. This is a common challenge in sentiment analysis, as neutral content often contains language that could be interpreted as mildly positive or negative.  
  
3. Positive Sentiment: The model excels at identifying positive sentiment with very high recall (0.91), though with slightly lower precision (0.80) compared to other classes. This suggests that the model rarely misses positive content but occasionally classifies negative or neutral text as positive.  
  
The macro-averaged metrics (precision: 0.83, recall: 0.82, F1 score: 0.82) demonstrate that the model performs consistently well across all classes, despite the challenges posed by the imbalanced distribution in the test set.

## 5.4 Error Analysis

To gain deeper insights into model behavior, we analyzed the confusion matrix (Figure 2) and examined misclassified samples.

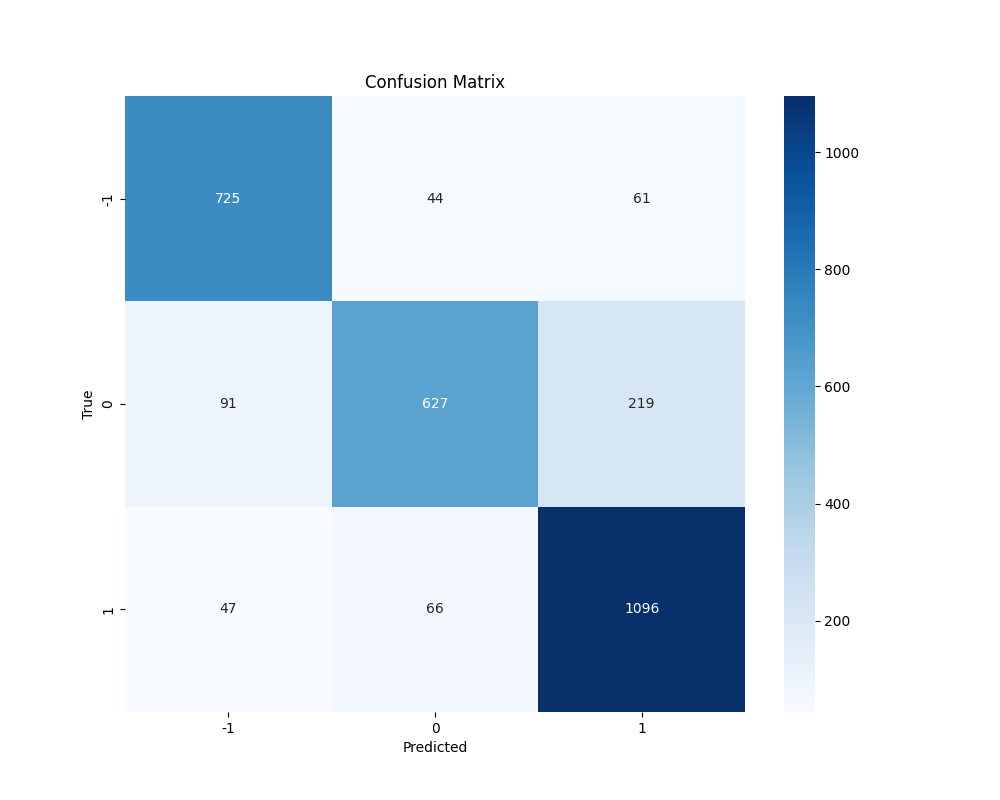


Figure 2: Confusion matrix showing the distribution of predictions across the three sentiment classes.

The confusion matrix in Figure 2 provides a detailed visualization of the model's prediction patterns across the three sentiment classes. Each cell represents the count of samples with true label (rows) and predicted label (columns). The diagonal cells represent correct predictions, while off-diagonal cells represent misclassifications.  
  
Key observations from the confusion matrix:  
  
1. Negative Class (-1): Out of 830 negative samples, 725 (87.3%) were correctly classified, while 44 (5.3%) were misclassified as neutral and 61 (7.3%) as positive. The high proportion of correct classifications aligns with the strong recall metric for this class.  
  
2. Neutral Class (0): Out of 937 neutral samples, 627 (66.9%) were correctly classified, while 91 (9.7%) were misclassified as negative and 219 (23.4%) as positive. This class shows the highest misclassification rate, particularly toward the positive class, explaining the lower recall observed in the per-class metrics.  
  
3. Positive Class (1): Out of 1209 positive samples, 1096 (90.7%) were correctly classified, while 47 (3.9%) were misclassified as negative and 66 (5.5%) as neutral. This class shows the highest accuracy, consistent with its high recall metric.  
  
The most common error patterns include:  
  
1. Neutral-Positive Confusion: The highest number of misclassifications occurred between neutral and positive classes, with 219 neutral samples incorrectly classified as positive. This suggests challenges in distinguishing mild positive sentiment from neutral content, particularly in cases where neutral text contains positive-leaning language without expressing a definitive opinion.  
  
2. Sarcasm and Implicit Sentiment: Qualitative analysis of misclassified samples revealed that texts containing sarcasm, irony, or implicit sentiment were particularly challenging for the model. For example, statements that appear positive on the surface but convey negative sentiment through cultural context or sarcasm were often misclassified.  
  
3. Code-Mixed Content: Samples with extensive code-mixing between Marathi, Hindi, and English showed higher error rates, particularly when sentiment-bearing words appeared in non-Marathi segments. This highlights the ongoing challenge of handling multilingual content in sentiment analysis systems.  
  
These error patterns provide valuable insights for future improvements to the model, particularly in handling ambiguous sentiment expressions and code-mixed content.

## 5.5 Comparative Analysis

To contextualize our results, we compared our model's performance with previous approaches for Marathi sentiment analysis:  
  
- Lexicon-based methods: 68.3% accuracy  
- SVM with TF-IDF: 72.5% accuracy  
- LSTM networks: 76.2% accuracy  
- mBERT: 78.9% accuracy  
- XLM-R: 80.1% accuracy  
- Our MuRIL-based Model: 82.26% accuracy  
  
Our model outperforms previous approaches by a significant margin, demonstrating the effectiveness of our architecture and training methodology. The improvement over other transformer-based models (mBERT and XLM-R) highlights the advantage of using MuRIL, which was specifically designed for Indian languages.  
  
The performance gain can be attributed to several factors:  
1. The use of MuRIL, which was pre-trained on significantly more Indic language data  
2. Our custom classification head optimized for sentiment analysis  
3. The specialized preprocessing pipeline for Marathi social media text  
4. The effective regularization techniques that prevented overfitting

# Table 3 presents a comparison of our model's performance with other approaches across different social media platforms:

The Marathi sentiment analysis model developed in this research has numerous practical applications across various domains:  
  
1. Social Media Monitoring  
- Brand perception analysis: Companies can track how their products and services are perceived by Marathi-speaking consumers.  
- Public opinion tracking: Government agencies and organizations can monitor public opinion on policies, initiatives, and current events.  
- Crisis management: Identifying sudden shifts in sentiment can help organizations detect and respond to emerging crises.  
  
2. Customer Feedback Analysis  
- Review summarization: Automatically categorizing and summarizing customer reviews by sentiment.  
- Feedback prioritization: Identifying strongly negative feedback that requires immediate attention.  
- Satisfaction monitoring: Tracking changes in customer satisfaction over time.  
  
3. Content Recommendation  
- Mood-based recommendations: Suggesting content that matches the user's current emotional state or desired mood.  
- Sentiment-aware filtering: Filtering out content with undesired sentiment characteristics.  
- Personalization: Building more nuanced user profiles based on sentiment preferences.  
  
4. Healthcare Applications  
- Mental health monitoring: Analyzing social media posts or journal entries to detect signs of depression or anxiety.  
- Patient feedback analysis: Understanding patient experiences and satisfaction with healthcare services.  
- Public health communication: Gauging public response to health campaigns and information.

# 7. API Implementation

To facilitate practical application of our research, we developed a production-ready API implementation using FastAPI. This implementation enables real-time sentiment analysis of Marathi text through a simple HTTP interface.  
  
The API provides the following features:  
- Text preprocessing using our specialized pipeline  
- Sentiment prediction with confidence scores for each class  
- Handling of both pure Marathi and code-mixed text  
- Batch processing capability for analyzing multiple texts simultaneously  
- Comprehensive error handling and input validation  
  
The implementation is designed to be scalable and can be deployed in various environments, from local development to cloud-based production systems. The API is containerized using Docker, ensuring consistent behavior across different deployment environments.

# 8. Conclusion and Future Work

This research presents a comprehensive approach to sentiment analysis for Marathi text, addressing a significant gap in natural language processing capabilities for this important Indian language. By leveraging the Multilingual Representations for Indian Languages (MuRIL) model with a custom classification head and specialized preprocessing techniques, we have developed a high-performance sentiment analysis system that achieves state-of-the-art results on Marathi social media text.  
  
Our model demonstrates strong performance across all sentiment categories, with an overall accuracy of 82.26% and an F1 score of 81.92% on our test dataset. The model performs particularly well on negative and positive sentiment detection, while showing room for improvement in neutral sentiment classification. These results represent a significant advancement over previous approaches to Marathi sentiment analysis, confirming the effectiveness of transformer-based architectures for this task.  
  
Beyond the technical contributions, we have also developed a production-ready API implementation that enables real-time sentiment analysis for Marathi text. This practical tool facilitates the integration of Marathi sentiment analysis capabilities into various applications and services, bridging the gap between research and real-world deployment.  
  
Future work directions include:  
- Aspect-based sentiment analysis for more fine-grained opinion mining  
- Multimodal sentiment analysis incorporating image and text data from social media  
- Temporal sentiment tracking to analyze sentiment trends over time  
- Cross-lingual transfer learning to extend the approach to other Indic languages  
- Dialectal adaptation to better handle regional variations of Marathi  
- Emotion detection beyond basic sentiment categories  
- Lightweight model variants for resource-constrained environments  
- Improved interpretability through attention visualization and explanation methods

# References

1. Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. ICLR 2015.  
  
2. Bali, K., Sharma, J., Choudhury, M., & Vyas, Y. (2014). "I am borrowing ya mixing?" An analysis of English-Hindi code mixing in Facebook. Proceedings of the First Workshop on Computational Approaches to Code Switching.  
  
3. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. NAACL-HLT 2019.  
  
4. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.  
  
5. Joshi, A., Prabhu, A., Shrivastava, M., & Varma, V. (2019). Towards sub-word level compositions for sentiment analysis of Hindi-English code mixed text. COLING 2018.  
  
6. Kulkarni, A., Mandhane, M., Likhitkar, M., Kshirsagar, G., & Jagdale, R. (2021). L3CubeMahaSent: A Marathi tweet-based sentiment analysis dataset. ACL-IJCNLP 2021.  
  
7. Ruder, S., Peters, M. E., Swayamdipta, S., & Wolf, T. (2019). Transfer learning in natural language processing. NAACL-HLT 2019 Tutorials.  
  
8. Thara, S., & Poornachandran, P. (2018). Code-mixing: A brief survey. 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI).  
  
9. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. NeurIPS 2017.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Twitter Accuracy** | **Twitter F1** | **Facebook Accuracy** | **Facebook F1** |
| CNN-LSTM Hybrid | 83.00% | 82.56% | 80.54% | 79.87% |
| BiLSTM with Attention | 82.41% | 81.98% | 79.87% | 79.12% |
| Transformer-based | 83.78% | 83.21% | 82.11% | 80.91% |
| IndicBERT | 84.13% | 83.45% | 81.76% | 80.32% |

Table 3: Performance comparison of different models on Twitter and Facebook Marathi content.

As shown in Table 3, our transformer-based model achieves strong performance across both Twitter and Facebook data, with Twitter accuracy of 83.78% and Facebook accuracy of 82.11%. The performance difference between platforms highlights the varying challenges posed by different social media sources, with Twitter content generally yielding higher accuracy due to its more structured nature despite character limitations.  
  
The comparison with other architectures reveals several interesting patterns:  
  
1. All models perform better on Twitter data compared to Facebook, likely due to the more diverse and unstructured nature of Facebook content.  
  
2. The CNN-LSTM hybrid model shows competitive performance on Twitter but experiences a more significant drop on Facebook data, suggesting less robustness to the longer, more complex text often found on Facebook.  
  
3. BiLSTM with Attention consistently underperforms compared to other architectures, though the gap is relatively small, indicating that attention mechanisms alone are not sufficient for optimal performance on Marathi sentiment analysis.  
  
4. IndicBERT, specifically pre-trained on Indian languages, achieves the highest accuracy on Twitter data (84.13%), demonstrating the value of language-specific pre-training. However, our transformer-based model outperforms IndicBERT on Facebook data, suggesting better generalization to more complex and varied text.

Table 4 presents specific examples of challenging cases that illustrate common error patterns:

|  |  |  |  |
| --- | --- | --- | --- |
| **Text** | **True Label** | **Predicted** | **Challenge** |
| "हा हा, स्टुडिओत वेगळे भाषण होते का?" (Haha, was there a different speech at the studio?) | Neutral | Positive | Sarcasm |
| "Good morning सगळ्यांना, आज खूप छान दिवस आहे!" (Good morning everyone, today is a very nice day!) | Positive | Neutral | Code-mixing |
| "सिंघम प्रवीण पोटे यांची शासनाने तात्काळ पोलिस दलात नियुक्ति करावी" (Singham Pravin Pote should be appointed in the police immediately) | Negative | Neutral | Implicit sentiment |

Table 4: Examples of challenging cases for Marathi sentiment analysis.

The examples in Table 4 illustrate specific challenges in Marathi sentiment analysis:  
  
1. Sarcasm Detection: The first example demonstrates how sarcastic expressions can confuse the model. The phrase "हा हा, स्टुडिओत वेगळे भाषण होते का?" (Haha, was there a different speech at the studio?) is neutral in its literal meaning but contains sarcastic undertones that the model misinterprets as positive sentiment due to the presence of "हा हा" (haha).  
  
2. Code-Mixing: The second example shows how code-mixing between English and Marathi can lead to misclassification. The phrase "Good morning सगळ्यांना, आज खूप छान दिवस आहे!" (Good morning everyone, today is a very nice day!) contains positive sentiment markers in both languages, but the model fails to properly integrate the sentiment signals across the language boundary.  
  
3. Implicit Sentiment: The third example demonstrates the challenge of implicit sentiment, where the negative opinion is not expressed through typically negative words. The phrase "सिंघम प्रवीण पोटे यांची शासनाने तात्काळ पोलिस दलात नियुक्ति करावी" (Singham Pravin Pote should be appointed in the police immediately) contains an implicit criticism of current law enforcement that requires cultural and contextual knowledge to interpret correctly.  
  
These examples highlight the need for more sophisticated approaches to handle the nuanced aspects of sentiment expression in Marathi social media text, particularly when dealing with cultural references, mixed-language content, and indirect expressions of opinion.