CAMERA: Context Based Emotion Detection Framework And Its Evaluation

Md Jahangir Alam[1], Ismail Hossain[1], Sai Puppala[2], and Sajedul Talukder[3]

[1] University of Texas at El Paso TX, USA
{malam10,ihossain}@miners.utep.edu
[2] Southern Illinois University Carbondale IL, USA
saimaniteja.puppala@siu.edu
[3] University of Texas at El Paso TX, USA
stalukder@utep.edu

Abstract. Detecting emotions online has increasingly become a critical concern for NLP researchers, particularly due to the proliferation of emotional expressions on social media and Web 2.0 platforms. Developing effective detection systems for languages with limited resources, such as Bengali, presents significant challenges. This paper introduces a novel dataset tailored specifically for contextual emotion detection in Bengali texts. The dataset creation process involved extensive data collection, preprocessing, human and automatic labeling, and label verification. This resulted in 20,247 annotated texts categorized into 27 different emotional categories. The dataset achieved a high Cohen's score of 0.89, indicating strong agreement among annotators. We define context with multiple components: 'WHO', 'WHEN', 'WHERE', and 'HOW'. Utilizing these context elements, we approximate a cognitive understanding of the posts, which facilitates emotion detection. We conducted comprehensive experiments using ML, DL, and BERT-based models to assess the dataset's efficacy. Our findings underscore the pivotal role of context in emotion detection. Particularly noteworthy was the performance of the BERT-based model XLM-R, which achieved an impressive F1 score of 0.88 and accuracy of 0.85 when context information was utilized. These results highlight how incorporating context significantly enhances the accuracy of emotion detection systems. This research contributes to advancing robust methodologies for identifying and understanding emotional content effectively.

Keywords: Bangla language · emotion detection · contextual emotion detection · social media post.

1 Introduction

Bangla, the fifth most common native Indo-European language and seventh globally, has about 300 million native and 37 million secondary speakers [8, 29].

Bangla, the primary language for 98% of Bangladeshis, is the national language of Bangladesh [5]. Its influence extends globally, with significant diasporas in the Middle East, Europe, and the USA [29].

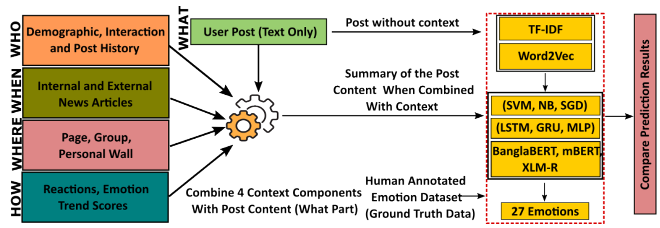

Fig. 1. Architecture of Emotion Detection With and Without Context

With the Digital Bangladesh initiative [21], Bangla's presence on platforms like Facebook, LinkedIn, and Twitter has expanded. Facebook, the dominant social platform in Bangladesh, has 33.71 million active Bengali users [22].

This manuscript introduces a new dataset and explores emotion detection techniques in Bengali social media text, specifically Facebook groups, pages, and individual profiles. Given the extensive research in English, Russian, and Arabic, the lack of work in Bangla highlights the importance of our study. Our work significantly advances Bengali emotion detection with a diverse emotion taxonomy, the concept of four context elements, and experiments using advanced language models.

Following are the main contributions of this paper:

- **Bangla Emotion Dataset Development:** We collect user posts and employ a semi-automated process to annotate the dataset.
- **Defining Context:** We define context with multiple components: 'WHO', 'WHEN', 'WHERE', and 'HOW'. These elements help approximate a cognitive understanding of the post, facilitating emotion detection.
- **Performance of Different Models on Emotion Detection:** We compare emotion classification experiments with and without context, using various ML, DL, and BERT-based models with different feature extraction methods.

2 Related Works

In 2020, Rayhan et al. [17] used a public Kaggle dataset translated into Bengali via Google Translator. The dataset contains 7,214 sentences, a vocabulary size of 57,000, and an embedding dimension of 64, with a maximum input length of 59. They employed two models for emotion classification: CNN-BiLSTM, achieving 66.62% accuracy, and BiGRU, achieving 64.96%. The study focused on six emotions: Happy, Fear, Sad, Angry, Surprise, and Love.

In 2020, Nath et al. [12] created a dataset from song lyrics for emotion classification. They experimented with various models, including Logistic Regression, SVM, Random Forest, Naive Bayes, LSVM, PSVM, KNN, and Decision Tree. Random Forest performed best, achieving 62% accuracy. The study classified lyrics into two emotions: Positive and Negative.

In 2020, Pran et al. [14] used 1,120 comments from coronavirus-related posts for emotion classification. They employed Word2Vec embeddings with CNN and LSTM models, with CNN achieving an impressive 97.24% accuracy. The study classified comments into three emotions: Analytical, Angry, and Depressed.

Lora et al. [9] used the Cricket and Restaurant datasets, comprising 4,468 rows with four columns. They utilized GloVe embeddings with models like CNN, RNN, Stacked LSTM, and Stacked LSTM with 1D convolution. The RNN model achieved the highest accuracy at 98%. The study focused on binary emotion classification: Positive and Negative.

In 2021, Purba et al. [15] collected 27,731 Bangla documents, annotated 995 samples, and categorized them into distinct emotions. Their models, including Logistic Regression, MNB, ANN, and CNN, achieved a peak accuracy of 68.27% using MNB. Emotions identified in their study encompassed Angry, Happy, and Sad.

Parvin et al. [13] gathered 8458 texts from various social media platforms like Facebook. Their models, employing Bag of Words (BoW) and Tf-idf representations, achieved a peak accuracy of 62% using SVM with Tf-idf. They categorized texts into six emotions: Anger, Fear, Disgust, Sadness, Surprise, and Joy.

In 2022, Rahib et al. [16] investigated emotion classification using 10,581 social media comments about COVID-19, annotated by the authors. Their models, such as SVM, Random Forest, CNN, and LSTM, achieved the highest accuracy of 84.92% with LSTM. The study concentrated on three emotions: Insightful, Curious, and Gratitude.

All these studies underscore the application of machine learning in understanding basic emotions across diverse linguistic and thematic contexts. No work incorporates context information like ours, and no one has introduced a larger taxonomy of emotions. We are the first to introduce the context with four components and larger number of emotion taxonomy.

3 Dataset

3.1 Context Data Collection

Demographic Data We collect Facebook users demographics such as name, work history, education, gender, and relationship status.

Interaction Data We compiled a dataset titled ‘User Interaction History in Social Networks with Other Users’ Posts’ focusing on activities such as commenting, sharing, and participating in discussions. This dataset offers insights into user behavior and social dynamics by employing a reverse engineering approach to gather user posts, metadata like usernames and profile URLs, and build an interconnected user network. This method reveals detailed interaction patterns, shedding light on individual behaviors and collective relationships within online communities. Upholding ethical standards and user privacy in accordance with platform policies and regulations, our approach not only captures interaction data but also facilitates deeper insights into social network structures, information diffusion, and community dynamics, benefiting sociological studies and social media content strategies.

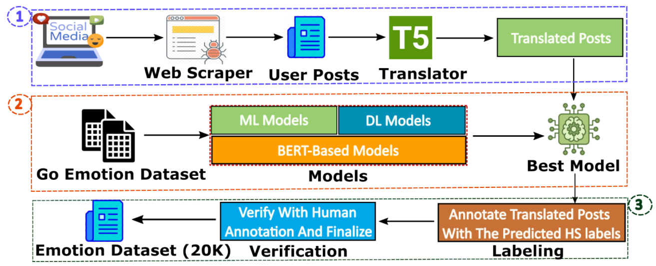

Fig. 2. Semi-automated Emotion Dataset Annotation Process.

3.2 User Post Collection

We curated a dataset from diverse Facebook groups and pages, encompassing a wide range of demographics, interests, and locations, totaling 20,247 posts. Our collection approach emphasized inclusivity and broad applicability, adhering to rigorous criteria. We gathered data efficiently while upholding ethical standards and respecting user privacy, aligned with Facebook's terms of service and data regulations.

3.3 Ethical Considerations

Our data collection followed ethical guidelines, securing permissions and consent from Facebook users and groups. We stored anonymized data and implemented measures to de-identify sensitive information. Approval from the Institutional Review Board was obtained to ensure compliance, with stringent measures in place to protect user privacy and uphold ethical standards throughout the study.

4 Methodology

In our research methodology, we first compile a dataset comprising Facebook posts from various sources such as groups, pages, and individual profiles. We annotate this data and perform semi-automated annotation, depicted in Figure 2. To enhance data quality, we apply diverse pre-processing algorithms to mitigate noise in the text. For evaluation, we train multiple machine learning models using our dataset, employing TF-IDF and word embedding techniques for feature extraction. We detail the model architectures tailored to these feature approaches. Finally, we conduct experiments on emotion detection, comparing results with and without context information.

4.1 Dataset Development Process

Our goal is to create a comprehensive Bengali emotion dataset with 27 distinct emotion categories and implement a context-based emotion detection technique. Inspired by the GoEmotions dataset developed by Google [3], we define these categories to classify emotional expressions written in Bengali. Developing such a dataset is challenging due to limited availability of suitable emotional text expressions, encountering issues like links, misspelled sentences, and 'Benglish' during data collection. Detecting emotions from plain text poses additional challenges compared to facial expression analysis, as text messages can mask underlying emotional states. Figure 2 outlines our dataset development process, encompassing data collection, preprocessing, human annotation, BERT-based automatic label prediction, and label verification phases, building upon methods described in [2].

Data Collection We implemented a semi-automatic annotation process for a dataset of 20,247 posts, following the approach in Ahmed et al. [28] and adhering to ethical and legal standards. We focused on Facebook groups, profiles, and pages with 1,000 to 5,000 members or followers. Our dataset included approximately 500 unique users to capture a range of Bangla language expressions. We ensured anonymity, complied with GDPR and local Bangladeshi regulations, and used cryptographic hash functions to de-identify personal information.

Preprocessing We addressed challenges in Bangla language processing for Facebook group discussions, where posts often include URLs, images, tags, and links. Our preprocessing removed non-Bangla characters, punctuation, URLs, images, links, hashtags, and user tags, aligning with previous studies [6, 7]. We refined a dataset of 20,247 posts into 27 emotion categories (excluding neutral) by managing stop words. We tokenized texts and eliminated Bangla stop words using a GitHub list [23], as referenced in [24]. This approach ensured clarity and reproducibility for accurate emotion detection.

Human Annotation The entire corpus underwent manual labeling, finalized through majority voting by a team of 5 graduate students specializing in Computer Science and conducting research in NLP.

Label Prediction In addition to human annotation, we employ automatic label prediction using the top-performing model selected from a range of ML, DL, and BERT-based models trained with the GoEmotions benchmark dataset. Among these models, mBERT demonstrates the highest performance [3]. Before making predictions, we translate Bengali text to English, as GoEmotions is an English dataset, and execute several preprocessing steps on the translated data for accurate predictions.

Label Verification We compare predicted labels with human-annotated labels. In cases of mismatches, posts are flagged for relabeling by human annotators.

The majority voting mechanism is used to determine human-annotated labels for mismatches. Instances where predicted labels match human-annotated labels less than 50% of the time prompt a second round of human annotation on more than half of the dataset.

4.2 Regular Emotion Detection

In this section, we outline our approach to emotion detection on our dataset, focusing solely on text without contextual elements. We employ TF-IDF and Word2Vec for feature extraction, as ML and DL models require processed inputs. We train and evaluate various ML, DL, and BERT-based models to compare their performance.

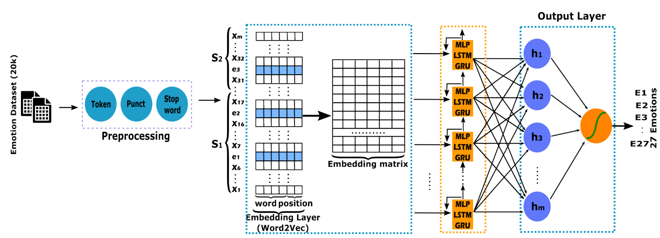

Fig. 3. Model Architecture for TF-IDF Feature.

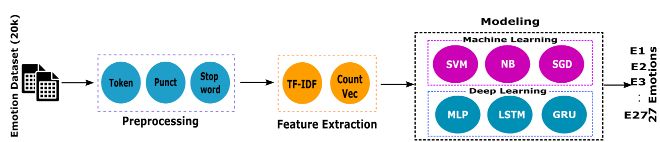

Fig. 4. Model Architecture for Word Embedding Feature.

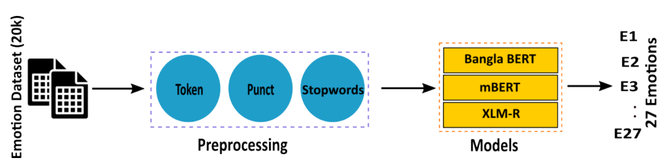

Fig. 5. Model Architecture for BERT-based Models.

Models with TF-IDF Features TF-IDF (term frequency-inverse document frequency) converts text into feature vectors by considering both term frequency (number of occurrences in a document) and inverse document frequency (number of documents containing the term) [10]. Implemented with `scikit-learn`, we focus on the top 1000 frequent words after preprocessing to remove non-Bangla characters. For emotion detection, we utilize traditional machine learning and deep learning models with TF-IDF features, incorporating n-grams (n=2) and a maximum of 1000 features. Figure 3 outlines these models. Additionally, we integrate Google's Word2Vec with CBOW and Skip-Gram architectures [11] to improve accuracy by capturing semantic relationships.

Models with Word Embedding Features Our approach preprocesses sentences by tokenizing them into 100-length one-hot encoding vectors using the top 1000 words. Longer sentences are truncated, and shorter ones are padded to fit this format. These vectors are then passed through an embedding layer initialized with Word2Vec embeddings, which can be trainable or fixed. The resulting 300-dimensional embeddings are processed through an LSTM layer for classification. The LSTM's output is fed into a dense layer for emotion detection, employing the Sigmoid activation function [25], RMSprop optimizer, and binary cross-entropy loss [19]. For an illustration of our model architecture using word embedding features, refer to Figure 4.

BERT-based Models We utilize three BERT-based models—Bangla-BERT, m-BERT, and XLM-R for classification tasks, leveraging their effectiveness and versatility [18, 4, 1]. These transformer models are fine-tuned using Huggingface and PyTorch.

Bangla BERT, based on the model ('sagorsarkar/Bangla-bert-base') described in [18], is pre-trained on a significant Bengali corpus. We conducted fine-tuning specific to our dataset, employing a batch size of 16 to optimize performance.

mBERT, introduced by [4], is pre-trained on a multilingual corpus, spanning 104 languages. We fine-tuned 'bert-base-multilingual-cased' model using a batch size of 16 for our study.

XLM-R, developed by [1], is a large-scale multilingual language model ('xlm-Roberta-base') trained across 100 diverse languages. We applied it to our dataset with a batch size of 16.

Each model underwent training for 5 epochs with a learning rate of 2e-5. The best-performing model based on validation results was selected for final predictions on the test dataset. For an illustration of our model architecture using BERT-based models, refer to Figure 5.

4.3 Contextual Emotion Detection

Figure 1 illustrates the process of emotion detection with and without context information. This process includes four contextual components: 'WHO', 'WHEN', 'WHERE', and 'HOW'. During standard emotion detection, only the text content of the user post is considered. In contrast, for contextual emotion detection,

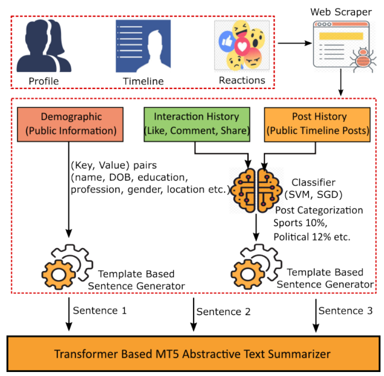

Fig. 6. Who component of the context

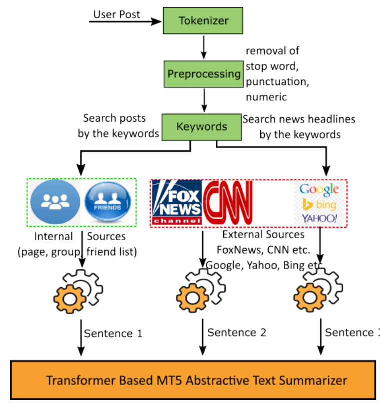

Fig. 7. When component of the context.

summaries of these four components are included along with the text content of the user post for classification. Subsequent sections detail the methodologies used for each of these contextual components.

4.4 "Who" Component Of The Context

In this section we describe how we perform the persona categorization, user post history categorization, user post interaction categorization. All these things comprise the who component of the context. Figure 6 depicts the flow of generating summary from the who component of the context.

User Persona Categorization In section, we gather a comprehensive user demographic dataset containing personal details such as name, age, gender, date of birth, education, profession, relationship status, and user interests. This dataset is annotated with 12 persona categories: political, businessperson, entrepreneur, educator, religious, artist, athlete, technologist, scientist, healthcare professional, traveler, and musician. Using machine learning models, we classify users into these persona categories, offering insights into their characteristics and traits to enhance contextual understanding.

User Post Categorization We collect user posts from their Facebook walls and utilize machine learning algorithms to classify these posts into 19 distinct categories. This categorization reveals the primary topics and themes prevalent in a user's posts. For example, one user may predominantly discuss politics, while another may focus on sports-related content. These insights illuminate user interests, preferences, and ideological tendencies, aiding in targeted content delivery and engagement strategies. Finally, we calculate the percentage distribution of each post category, such as 10% political posts, 12% sports posts, 15% religious posts, and so on.

User Interaction Post Categorization Within the Facebook ecosystem, users actively engage with posts from friends, groups, and pages. We curate a list of these interactions and annotate them accordingly. Using machine learning algorithms, we categorize these interaction posts into 19 distinct categories. This categorization, akin to user post analysis, reveals patterns in user engagement, showcasing preferences for specific topics, interests, or ideologies. Understanding these dynamics facilitates personalized content recommendations and targeted engagement strategies. We also calculate the percentage distribution of each category, similar to how we analyze user posts.

4.5 "When" Component Of The Context

Our methodology for temporal analysis of social media content employs advanced Natural Language Processing (NLP) techniques, focusing particularly on analyzing the temporal aspect of user-generated text. This approach encompasses several key steps detailed below.

The user's post is tokenized to break it down into individual words or tokens. This process involves removing stop words, punctuation marks, and non-alphanumeric characters, retaining only meaningful content. Following preprocessing, significant keywords are extracted from the post. These keywords serve as crucial indicators of the central themes or topics within the user's message. Techniques such as TF-IDF (Term Frequency-Inverse Document Frequency) or other NLP algorithms are utilized to identify and rank the importance of these keywords. Using the extracted keywords, searches are conducted within the Facebook platform to retrieve contemporary news articles or posts closely related to the identified topics. This includes searching posts from the user's friend list, Facebook pages, and public Facebook groups. Similarly, the identified keywords are used to initiate searches on the Google search engine. This retrieves news articles, blog posts, or other web content across the internet that align with the topics identified in the user's post. This process involves developing a Python script to perform keyword-based searches on Google. The retrieved news articles and posts from both Facebook and Google are aggregated and analyzed. This provides a comprehensive overview of contemporary events and topics that correspond with the themes found in the user's post.

This approach enables us to gain insights into temporal trends and current events relevant to user-generated content, enhancing our understanding of user interests and engagement patterns.

4.6 "Where" Component of the Context

We compile a comprehensive user post dataset using a diverse data collection strategy that gathers posts from various sources within the Facebook ecosystem. This includes aggregating posts from Facebook pages, user profiles, groups, and other relevant sources. By systematically collecting data from these channels, we capture a wide range of user-generated content, encompassing diverse perspectives, topics, and levels of engagement. This extensive dataset serves as a valuable resource for our research and analysis, providing insights into user behavior and interaction dynamics on Facebook. Figure 8 illustrates the architecture used for collecting information regarding the "where" component of the context.

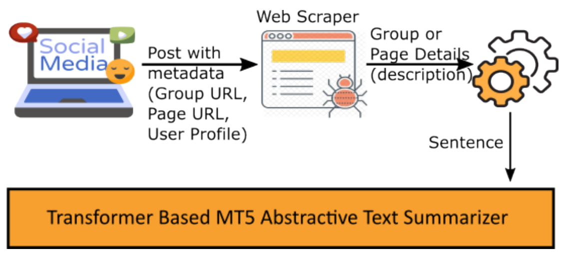

Fig. 8. Where component of the context.

4.7 "How" Component of the Context

In our study [7] accepted for ICWSM 2024, we focus on analyzing user posts' impact on audience engagement across social networks. We collect and analyze user interactions such as likes, comments, and shares, using sentiment trend analysis to understand the emotional responses within comment sections. This approach helps assess how users influence audience sentiment and provides insights into online discourse dynamics and user interactions on social media platforms. As shown in Figure 9 through our methodology, we aim to distill these insights into clear understandings of user engagement dynamics, integrating them with contextual elements like who, when, and where for comprehensive analysis.

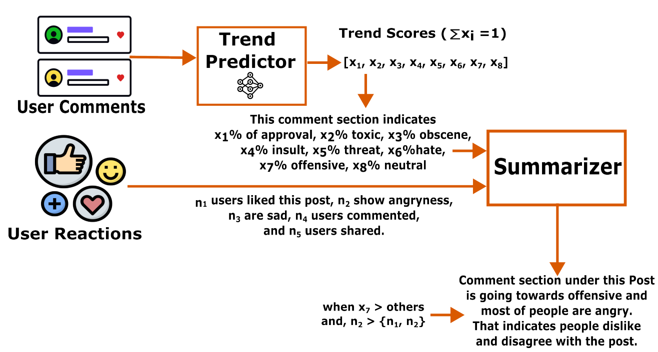

Fig. 9. How Component of the Context

Finally combining all the 4 components of the context we pass it to summarizer. Along with the summary of the 4 components we pass the text of the original post to the emotion classifier as shown in Figure 1, which is contextual emotion detection.

5 Experiment

In this section we discuss the experimental details of the context components, experimental details of regular and contextual emotion detection.

5.1 Who Part Experiment

In this section we describe experiment related to who part of the context object. We experiment on user persona categorization, user post categorization, and user interaction post categorization. We categorize posts and calculate the percentage of post categories.

We analyzed user demographics, posts, and interactions on Facebook to understand behavior and engagement patterns. Demographic data from 500 users included personal details, education, profession, and page likes. Using SVM, NB, and SGD models, users were categorized into 12 personas. From 20,247 user posts, categorized into 19 types, the best-performing SVM model predicted category percentages for each user. We used similar models to categorize user interaction posts for contextual emotion detection.

5.2 When Part Experiment

In this section, we outline our experimental approach for collecting contemporary news and articles from both internal (Facebook ecosystem) and external (search engines) sources.

Contemporary Events News And Articles From Internal and External Sources We streamline the analysis of user-generated content on social media by extracting relevant articles from both internal and external sources. First, we tokenize the text. Next, we preprocess it by removing stop words, numbers, and punctuation. Using GPT and NLP techniques, we then extract significant keywords. These keywords guide searches through users' Facebook friends' groups and lists to identify posts relevant to the research criteria.

For external sources, the process involves tokenization, preprocessing, and keyword extraction. Keywords are then used to search platforms like news websites and search engines (Google, Bing, Yahoo), ensuring efficient retrieval of relevant articles aligned with research objectives. Figure 7 illustrates this workflow.

Summarize The Contemporary Events News And Articles Using GPT4
We use GPT-4 and prompt engineering via the completion API to generate concise summaries of news articles. These summaries are then amalgamated into a comprehensive overview using GPT-4, providing nuanced insights into global events. This synthesized text enriches our analysis by embedding findings within a broader narrative framework, enhancing relevance and contextual emotion detection.

5.3 Where Part Experiment

We capture metadata about where users post within the Facebook ecosystem, including their personal walls, groups, or pages. This contextual information helps understand where and with whom users interact, which can influence contextual emotion detection in their posts.

5.4 How Part Experiment

We fine-tuned RoBERTa on our dataset [7], achieving an F1-score of 0.75 and an AUC of 0.92. The RoBERTa outputs were used to train XGBoost, which achieved an F1-score of 0.79, outperforming other models. XGBoost predicted trend scores based on RoBERTa's emotion outputs for 28 emotions. Comments were analyzed by RoBERTa, aggregated at the reply tree level, and processed by XGBoost to generate final trend probability scores for contextual emotion detection.

6 Results

In this section we discuss the experimental results of contextual components which are post categorization, persona categorization, interaction post categorization.

6.1 Performance of Post Categorization

We trained SVM, Naive Bayes (NB), and Stochastic Gradient Descent (SGD) models to categorize user posts. SGD emerged as the best performing model with an accuracy of 0.88, as shown in Table 1.

6.2 Performance of Persona Categorization

We trained machine learning models to categorize user personas into types such as political, athletic, business, and religious. Table 2 summarizes their performance. The NB model achieved the highest accuracy of 0.82, making it our selected model for persona categorization.

Model	Accuracy	F1
SVM	0.85	0.85
NB	0.80	0.80
SGD	**0.88**	**0.85**

Table 1. Performance Comparison of ML Models in Post Categorization

Model	Accuracy	F1
SVM	0.78	0.78
NB	**0.82**	**0.82**
SGD	0.80	0.81

Table 2. Performance Comparison of ML Models in Persona Categorization

Hyperparameter	M1	M2	M3
Learning rate	2e-5	2e-5	2e-5
Epochs	5	5	5
Batch size	16	16	16
Dropout	0.2	0.2	0.2
Max seq length	128	128	128

Table 3. List of hyperparameters for BERT-based models. Here M1 = Bangla BERT, M2 = mBERT, M3 = XLM-R.

Ref.	Acc. (%)	No. of Emotions	Data Size
[17]	66.62	6	7k
[16]	84.92	3	10k
[15]	68.27	3	27k
[13]	62	6	8k
[17]	62.62	6	7k
[2]	69.61	6	6k
This paper	**85**	**27**	20k

Table 4. Comparison with existing work.

7 Testing in the Wild: Comparison Between Regular and Contextual Emotion Detection

In this experiment, we compare regular and contextual emotion detection using 100 user posts annotated by multiple interpreters. Discrepancies were resolved through discussions to ensure accurate ground truth data. This benchmark was then used to evaluate various ML, DL, and BERT-based models.

Table 3 details the hyperparameters for training BERT-based models on our dataset. Table 5 presents performance metrics for emotion detection with and without context using various ML and DL models. Among ML models, SVM and NB have lower accuracies compared to SGD, with SVM performing the worst with context (accuracy: 0.60, F1: 0.62). For DL models, the MLP performs the worst with context (accuracy: 0.41, F1: 0.45). XLM-R is the top performer among BERT-based models with context (accuracy: 0.84, F1: 0.88). BERT-based models generally outperform traditional ML and DL models in both scenarios, demonstrating their effectiveness in handling contextual information for emotion detection tasks.

7.1 Comparison with Existing Work

Recent advancements in emotion detection research have explored various methodologies, from traditional machine learning to sophisticated deep learning approaches. Table 4 compares these studies, highlighting dataset choices and evaluation methods. While most studies use a standard set of six basic emotion categories, our work introduces a broader taxonomy of 27 emotion categories and integrates four comprehensive contextual elements for enhanced emotion detection.

8 Ethical Considerations

We've created an automated script to ethically collect and analyze social media data under IRB approval. Our analysis uses only anonymized data collected from various sources (groups, pages, profiles), which does not meet the PII criteria defined in NIST SP 800-122 [27]. According to GDPR guidelines [26], information used without context, such as names or personal identification numbers, does not constitute "personal information".

9 Discussion and Limitations

The dataset of 100 user posts, annotated with human judgments, guided our evaluation of regular versus contextual emotion detection. However, its small size may limit generalizability. Future research should use larger and more diverse datasets from various social media platforms for broader insights.

Our method integrating contextual information has shown improvements, yet its effectiveness relies on accurately extracted contextual data, which can be sparse or noisy. Enhancing context extraction and noise reduction methods could enhance robustness.

While advanced models like XLM-R excel, they are computationally intensive and may struggle with highly specialized content. Future studies should explore domain-specific models and continual learning to address these challenges.

Models	**With Context**		**Without Context**	
	Accuracy	**F1**	**Accuracy**	**F1**
ML Models				
SVM	0.60	0.62	0.62	0.67
NB	0.61	0.66	0.63	0.68
SGD	0.70	0.76	0.68	0.72
DL Models				
LSTM	0.64	0.69	0.62	0.68
GRU	0.62	0.66	0.43	0.47
MLP	0.41	0.45	0.44	0.47
BERT-based Models				
Bangla BERT	0.79	0.81	0.67	0.72
mBERT	0.79	0.83	0.75	0.78
XLM-R	**0.85**	**0.88**	**0.79**	**0.82**

Table 5. Performance Metrics for Emotion Detection with and without Context

10 Conclusion & Future Work

This paper introduces a novel approach to emotion detection by integrating contextual information through advanced models such as Bangla BERT, mBERT, and XLM-R. Our results show that contextual emotion detection surpasses regular methods, achieving higher accuracy scores and closer alignment with human judgments. These findings highlight the potential of advanced AI models to improve emotion detection accuracy by emulating human cognitive processes. Future research will explore additional modalities, refine integration techniques, and explore broader implications for social media analytics and user experience design.

11 Acknowledgment

This research was supported by NSF grant CNS-2153482.

References

1. Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., Stoyanov, V.: Unsupervised cross-lingual representation learning at scale. arXiv preprint arXiv:1911.02116 (2019)
2. Das, A., Sharif, O., Hoque, M. M., Sarker, I. H.: Emotion classification in a resource constrained language using transformer-based approach. arXiv preprint arXiv:2104.08613 (2021)
3. Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., Ravi, S.: GoEmotions: A dataset of fine-grained emotions. arXiv preprint arXiv:2005.00547 (2020)
4. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
5. Eglitis-media.: worlddata.info. (2024). https://www.worlddata.info/languages/bengali.php. Accessed 2 August 2024
6. Hossain, I., Puppala, S., Alam, Md J., Talukder, S.: Monitoring Dynamics of Emotional Sentiment in Social Network Commentaries. In: Proceedings of the International Conference on Advances in Social Networks Analysis and Mining, pp. 51–55 (2023)
7. Hossain, I., Puppala, S., Alam, Md J., Talukder, S.: A Visual Approach to Tracking Emotional Sentiment Dynamics in Social Network Commentaries (2024)
8. Klaiman, M. H., Lahiri, A.: Bengali. In: The world's major languages, pp. 427–446. Routledge (2018)
9. Lora, S. K., Jahan, N., Antora, S. A., Sakib, N.: Detecting emotion of users' analyzing social media bengali comments using deep learning techniques. In: 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), pp. 88–93. IEEE (2020)
10. Luthfi Ramadhan.: TF-IDF Simplified. (2021). https://towardsdatascience.com/tf-idf-simplified-aba19d5f5530. Accessed 2 August 2024
11. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781 (2013)
12. Nath, D., Roy, A., Shaw, S. K., Ghorai, A., Phani, S.: Textual lyrics based emotion analysis of bengali songs. In: 2020 International Conference on Data Mining Workshops (ICDMW), pp. 39–44. IEEE (2020)
13. Parvin, T., Hoque, M. M.: An ensemble technique to classify multi-class textual emotion. Procedia Computer Science, vol. 193, pp. 72–81. Elsevier (2021)

14. Pran, Md S. A., Bhuiyan, Md R., Hossain, S. A., Abujar, S.: Analysis of Bangladeshi people's emotion during COVID-19 in social media using deep learning. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1–6. IEEE (2020)
15. Purba, S. A., Tasnim, S., Jabin, M., Hossen, T., Hasan, Md K.: Document level emotion detection from bangla text using machine learning techniques. In: 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), pp. 406–411. IEEE (2021)
16. Rahib, Md R. H. K., Tamim, A. H., Tahmeed, M. Z., Hossain, M. J.: Emotion detection based on Bangladeshi people's social media response on COVID-19. SN Computer Science, vol. 3, no. 2, p. 180. Springer (2022)
17. Rayhan, Md M., Al Musabe, T., Islam, Md A.: Multilabel emotion detection from bangla text using bigru and cnn-bilstm. In: 2020 23rd International Conference on Computer and Information Technology (ICCIT), pp. 1–6. IEEE (2020)
18. Sagor Sarker.: BanglaBERT: Bengali Mask Language Model for Bengali Language Understanding. (2020). https://github.com/sagorbrur/bangla-bert
19. Saxsena, S.: Binary Cross Entropy/Log Loss for Binary Classification. Log Loss for Binary Classification, pp. 02–08 (2021)
20. Sharma, O.: A New Activation Function for Deep Neural Network. In: International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), pp. Feb 14-16, Faridabad, India (2019)
21. Simon Kemp.: DIGITAL 2021: BANGLADESH. (2021). https://datareportal.com/reports/digital-2021-bangladesh. Accessed 2 August 2024
22. StatCounter Global Stats.: Social Media Stats in Bangladesh. (2024). https://gs.statcounter.com/social-media-stats/all/bangladesh. Accessed 2 August 2024
23. stopwords-iso.: Stopwords Bengali. (2024). https://github.com/stopwords-iso/stopwords-bn. Accessed 2 August 2024
24. Tripto, N. I., Ali, M. E.: Detecting Multilabel Sentiment and Emotions from Bangla YouTube Comments. In: International Conference on Bangla Speech and Language Processing (ICBSLP), pp. Sept 21-22, Sylhet, Bangladesh (2018)
25. Sharma, O.: A new activation function for deep neural network. In: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), pp. 84–86 (2019)
26. General Data Protection Regulation (GDPR). In: GDPR Info (2021). https://gdpr-info.eu/. Accessed 2 August 2024
27. Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). In: NIST (2021). https://tinyurl.com/ylyjst5y. Accessed 2 August 2024
28. Ahmed, S., Alam, Md J., Talukder, S., Hossain, I.: Towards Addressing Identity Deception in Social Media using Bangla Text-Based Gender Identification. In: Proceedings of the International Conference on Advances in Social Networks Analysis and Mining, pp. 72–76 (2023)
29. Chung Hwan Kwak.: New World Encyclopedia. (2020). https://www.newworldencyclopedia.org/entry/Bengali_language. Accessed 2 August 2024